
UNIVERSITY OF TORONTO
3 1761 01027892 7

GRUNDZÜGE

DER

MENGENLEHRE

VON

FELIX HAUSDORFF
O. PROFESSOR DER MATHEMATIK AN DER UNIVERSITÄT GREIFSWALD

MIT 53 FIGUREN IM TEXT



LEIPZIG
VERLAG VON VEIT & COMP.
1914



Druck von Metzger & Wittig in Leipzig.

DEM SCHÖPFER DER MENGENLEHRE

HERRN GEORG CANTOR

IN DANKBARER VEREHRUNG

GEWIDMET

Vorwort.

Das vorliegende Werk will ein Lehrbuch und kein Bericht sein: es versucht die Hauptsachen der Mengenlehre ohne Voraussetzung höherer Vorkenntnisse mit vollständig ausgeführten Beweisen darzustellen und verzichtet dafür auf Vollständigkeit des behandelten Stoffes. Hinsichtlich der thematischen Begrenzung, bei der ja übrigens die Mitwirkung subjektiver Gründe nicht auszuschalten ist, habe ich der Mengenlehre selbst vor ihren Anwendungen den Vorzug gegeben; infolgedessen wird man vielleicht finden, daß den geordneten Mengen zuviel und etwa den reellen Funktionen einer reellen Variablen zuwenig Platz eingeräumt worden sei. Was die Tendenz anbelangt, immer zu beweisen und niemals bloß zu referieren, so ist mir die in dem bekannten Voltaireschen Worte bezeichnete Gefahr nicht entgangen; aber in einem Gebiet, wo schlechthin nichts selbstverständlich und das Richtige häufig paradox, das Plausible falsch ist, gibt es außer der lückenlosen Deduktion kaum ein Mittel, sich und den Leser vor Täuschungen zu bewahren. Ich habe, um von dem menschlichen Privileg des Irrtums einen möglichst sparsamen Gebrauch zu machen, nichts ungeprüft übernommen und manches von der Wiedergabe ausgeschlossen, was mir nur auf persönlichen Kredit hin glaubwürdig erschien; aber selbst fertige und im ganzen einwandfreie Darstellungen, die ich meinem Plane einzugliedern hatte, mußte ich häufig einer gründlichen Umarbeitung unterziehen, bis sie sich den mir vorschwebenden Forderungen an Präzision fügten.

Nach diesem Programm glaube ich, daß das Buch von jedem, der über einige Abstraktion des Denkens verfügt, insbesondere oder außerdem von Studierenden der Mathematik in mittleren Semestern mit Erfolg gelesen werden kann; auf der andern Seite würde ich seinen Zweck für verfehlt halten, wenn ich nicht hoffen dürfte, auch den Fachgenossen manches Neue, mindestens in methodischer

und formaler Hinsicht, zu bieten. Daß man zerstreute Einzelheiten logisch verkettet und systematisiert, bisherige Resultate von unnötig speziellen und komplizierenden Voraussetzungen befreit, einen Fortschritt in Einfachheit und Allgemeinheit anstrebt, ist schließlich das mindeste, was von dem Bearbeiter eines schon behandelten Stoffes verlangt werden kann. Diesen Anforderungen hoffe ich innerhalb gewisser Grenzen entsprochen und dem Gegenstande wenigstens einige neue Seiten abgewonnen zu haben; um zur Rechtfertigung dieses Anspruches nicht nur auf das Buch als Ganzes zu verweisen, mögen hier etwa die symmetrischen Mengen, die Limesbildungen von Mengenfolgen, die Umgebungstheorie der Punktmengen, die systematische Durchführung der Relativbegriffe, die reduziblen Mengen, die Behandlung des Inhalts von Punktmengen erwähnt werden.

Die Quellenangaben sind in einen Anhang verwiesen, der außerdem einige nicht unwesentliche Nachträge bringt. Ich glaubte jene in Grenzen halten zu dürfen, die man in unserem historisch-philologischen Zeitalter entschieden eng finden wird, und zwar nicht nur aus äußeren Gründen, sondern weil auch eine innere Schwierigkeit bestand, Keime und Anregungen, die sich im Laufe der Darstellung bisweilen erheblich umgeformt hatten, nachträglich zu rekonstruieren. Insbesondere sind durch die Axiomatisierung der Punktmengentheorie viele Sätze über lineare Punktmengen so verwandelt, verallgemeinert, zerlegt und in einem andern Zusammenhang wieder verknüpft worden, daß ein einfaches Zitat kein richtiges Bild geben kann.

Herr J. O. Müller (Bonn) hat sich der aufopferungsvollen Mühe unterzogen, eine Korrektur des Buches mitzulesen, und Herr W. Blaschke (Prag) den größten Teil der Figuren gezeichnet; beiden Kollegen fühle ich mich zu herzlichem Danke verpflichtet.

Greifswald, 15. März 1914.

Felix Hausdorff.

Inhalt.

Erstes Kapitel.

Mengen und ihre Verknüpfungen: Summe, Durchschnitt, Differenz.

§ 1. Der Mengenbegriff.

Eine Menge ist eine Zusammenfassung von Dingen zu einem Ganzen, d. h. zu einem neuen Ding. Man wird dies allerdings schwerlich als Definition, sondern nur als anschauliche Demonstration des Mengenbegriffs gelten lassen, die auf einfache Beispiele verweist: wie etwa die Menge der Einwohner einer Stadt, die Menge der Wasserstoffatome in der Sonne. Diese beiden Mengen sind endlich, sie bestehen aus einer endlichen, die zweite freilich aus einer ungeheuer großen Anzahl von Gegenständen. Es ist das Verdienst Georg Cantors, auch unendliche, d. h. nicht endliche Mengen in den Kreis der Betrachtung gezogen und damit, über populäre Vorurteile und philosophische Machtsprüche hinwegschreitend, eine neue Wissenschaft, die Mengenlehre, begründet zu haben; denn eine bloße Theorie der endlichen Mengen wäre ja nichts weiter als Arithmetik und Kombinatorik. Die Menge der natürlichen Zahlen, die Menge der Punkte des Raumes sind die nächstliegenden Beispiele unendlicher Mengen.

Die Mengenlehre ist das Fundament der gesamten Mathematik; Differential- und Integralrechnung, Analysis und Geometrie arbeiten in Wirklichkeit, wenn auch vielleicht in verschleiernder Ausdrucksweise, beständig mit unendlichen Mengen. Über das Fundament dieses Fundamentes, also über eine einwandfreie Grundlegung der Mengenlehre selbst ist eine vollkommene Einigung noch nicht erzielt worden. Die nächstliegenden Schwierigkeiten und Vorurteile dürfen zwar als erledigt gelten: viele anscheinende „Paradoxien des Unendlichen" sind nur so lange paradox, wie man an der unberechtigten Forderung festhält, daß für endliche und für unendliche Mengen unterschiedslos dieselben Gesetze gelten sollen. Die naturgemäßen Abweichungen zwischen beiden Gebieten bedingen keinen Widerspruch innerhalb des Unendlichen. Dagegen ist eine wirkliche Paradoxie noch nicht befriedigend aufgeklärt, auf die der naive Mengen-

begriff, mit seiner Zusammenfassung beliebig vieler Elemente zu einer Menge, letzten Endes hinausführt. In einer typischen Form, die allerdings noch der mathematischen Bestimmtheit entbehrt, lautet dies Paradoxon folgendermaßen: wenn es zu jeder Menge von Dingen noch ein weiteres, von ihnen allen verschiedenes Ding gibt, so ist die Gesamtheit aller Dinge offenbar selbst keine Menge. Ein solches System von Dingen, das nicht als Menge aufgefaßt werden kann, bilden, wie wir sehen werden, die (endlichen und unendlichen) Kardinalzahlen oder auch die Ordnungszahlen. Den hiernach notwendigen Versuch, den Prozeß der uferlosen Mengenbildung durch geeignete Forderungen einzuschränken, hat E. Zermelo unternommen. Da indessen diese äußerst scharfsinnigen Untersuchungen noch nicht als abgeschlossen gelten können und da eine Einführung des Anfängers in die Mengenlehre auf diesem Wege mit großen Schwierigkeiten verbunden sein dürfte, so wollen wir hier den naiven Mengenbegriff zulassen, dabei aber tatsächlich die Beschränkungen innehalten, die den Weg zu jenem Paradoxon abschneiden.

Wir fassen also irgendwelche Dinge $a, b, \ldots$, endlich oder unendlich viele, zu einer Menge A zusammen; diese Dinge heißen die Elemente der Menge A. Wir sagen auch, a gehört zu A, a liegt in A, a ist in A enthalten, oder A hat das Element a, A enthält a. Von der Beschaffenheit dieser Elemente sieht die reine Mengenlehre gänzlich ab; indessen empfehlen wir dem Leser, sich nach Möglichkeit mathematische Objekte (Punkte, Figuren, Zahlen, Funktionen u. dgl.) darunter vorzustellen.

Zwei Mengen A, B werden dann und nur dann als gleich betrachtet ($A = B$), wenn sie genau dieselben Elemente enthalten, wenn also jedes Element der einen auch Element der andern ist. Eine solche Gleichung $A = B$ kann, wenn sie zwei formal verschieden definierte Mengen identifiziert, recht wohl bedeutenden Erkenntniswert besitzen: wenn man will, könnte man jede noch so tief liegende mathematische Wahrheit in das Gewand einer Mengengleichung kleiden. Z. B. läßt sich der berühmte letzte Fermatsche Satz, mag er richtig sein oder nicht, jedenfalls durch die Behauptung ersetzen: die Menge der Primzahlen p, für welche die Gleichung $x^p + y^p = z^p$ in natürlichen Zahlen x, y, z auflösbar ist, ist identisch mit der Menge der geraden Primzahlen (beide Mengen bestehen aus der einzigen Zahl 2).

Die aus den Elementen $a, b, \ldots$ bestehende Menge wird mit

$$A = \{a, b, \ldots\}$$

bezeichnet; die Punkte deuten das Vorhandensein weiterer Elemente an. Die Menge $\{a, b\}$ besteht aus den zwei Elementen a, b. Die Menge $\{a\}$, die aus dem einzigen Element a besteht, ist von diesem

Element selber zunächst sorgfältig zu unterscheiden, schon aus dem Grunde, weil wir auch Mengen (Systeme) zulassen wollen und müssen, deren Elemente selbst wieder Mengen sind. Die Menge $a = \{1, 2\}$ besteht aus den zwei Elementen 1, 2, die Menge $\{a\}$ aus dem einzigen Element a. In vielen Fällen ist allerdings die Verschmelzung beider Dinge ganz unbedenklich.

Außer den Mengen, die Elemente haben, lassen wir auch eine Menge 0, die Nullmenge, zu, die kein Element hat; die Gleichung $A = 0$ bedeutet also, daß auch die Menge A kein Element hat, verschwindet, leer ist.[1] Auch hierzu ist die analoge Bemerkung zu machen, wie zum allgemeinen Fall: die Gleichung $A = 0$ kann eine bedeutungsvolle Aussage sein, wenn nämlich die Definition der Menge A ihr Verschwinden nicht unmittelbar erkennen läßt. Der Fermatsche Satz behauptet: die Menge der natürlichen Zahlen $n > 2$, für welche die Gleichung $x^n + y^n = z^n$ in natürlichen Zahlen x, y, z lösbar ist, ist die Nullmenge.

§ 2. Teilmengen, Differenzen.

Wenn alle Elemente der Menge A auch Elemente der Menge B sind, so sagen wir: A ist in B enthalten[2], A ist eine Teilmenge von B, eine Menge unter B, oder B enthält A, B ist eine Menge über A. Wir bringen dies durch eine der beiden Formeln

$$A \leqq B \quad \text{oder} \quad B \geqq A$$

zum Ausdruck; wobei die Zeichen $\leqq$ $\geqq$ an die üblichen Zeichen $<$ $>$ für kleiner und größer erinnern, aber doch von ihnen unterschieden werden sollen. Zu den Teilmengen von B rechnen wir auch die Menge B selbst und die Nullmenge: eine wichtige Verabredung, deren Zweckmäßigkeit sich vielfach bewähren wird.

Die Teilmengen der aus 4 Elementen bestehenden Menge $\{a, b, c, d\}$ sind:

$$0$$
$$\{a\} \quad \{b\} \quad \{c\} \quad \{d\}$$
$$\{a, b\} \quad \{a, c\} \quad \{a, d\} \quad \{b, c\} \quad \{b, d\} \quad \{c, d\}$$
$$\{a, b, c\} \quad \{a, b, d\} \quad \{a, c, d\} \quad \{b, c, d\}$$
$$\{a, b, c, d\}$$

Ihre Anzahl ist $1 + 4 + 6 + 4 + 1 = 16 = 2^4$.

---

[1] Die Ausdrucksweise: „die Menge A existiert nicht", können wir nicht akzeptieren. Sie existiert, aber es existieren keine Elemente von ihr.

[2] Es ist die Forderung erhoben worden, die Beziehung einer Menge zu ihren Teilmengen und zu ihren Elementen verschieden auszudrücken, z. B. eine Menge enthält ihre Teilmengen, besitzt ihre Elemente, oder sie umfaßt ihre Teilmengen, enthält ihre Elemente. Wir halten eine solche Fesselung der Sprache für schwer durchführbar und für überflüssig.

Die Anzahl der Teilmengen mit m Elementen einer Menge mit n Elementen ist der Binomialkoeffizient

$$\binom{n}{m} = \frac{n(n-1)(n-2)\ldots(n+1-m)}{1.\quad 2.\qquad 3.\quad \ldots \qquad m}, \quad \left(\binom{n}{0} = 1\right)$$

und die Anzahl aller Teilmengen einer Menge mit n Elementen ist

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = (1+1)^n = 2^n.$$

Wenn zwei Mengen gleichzeitig Teilmengen voneinander sind, so sind sie identisch, d. h. aus

$$A \leqq B, \quad B \leqq A$$

folgt

$$A = B.$$

In dieser Gestalt wird sehr häufig die Identität zweier Mengen bewiesen werden.

Wenn A Teilmenge von B, nicht aber B Teilmenge von A, die Gleichheit beider Mengen also ausgeschlossen ist, so schreiben wir auch

$$A < B \quad \text{oder} \quad B > A;$$

in diesem Falle wird A gelegentlich als echte Teilmenge von B bezeichnet werden. B enthält dann also mindestens ein Element, das nicht auch Element von A ist. Diese Bezeichnung erlaubt auch den Tatbestand auszudrücken, daß eine Menge B nicht Null sein, also mindestens ein Element besitzen soll, nämlich durch

$$0 < B \quad \text{oder} \quad B > 0.$$

Wenn $A \leqq B$, so ist entweder $A < B$ oder $A = B$.

Wenn $A < B$, $B < C$, so ist auch $A < C$; man sagt, daß die durch das Zeichen $<$ ausgedrückte Relation das transitive Gesetz befolge. Gleiches gilt von dem Zeichen $>$. Wenn $A \leqq B \leqq C$, so nennen wir B gelegentlich eine Menge zwischen A und C.

Wenn A Teilmenge von B ist, so verstehen wir unter

$$B - A$$

die Menge derjenigen Elemente von B, die nicht Elemente von A sind; wir nennen diese Menge das Komplement von A in B und haben damit also die Differenz zweier Mengen definiert, unter der Voraussetzung, daß der Subtrahend Teilmenge des Minuenden ist. Diese Menge $B - A$ ist ihrerseits eine Teilmenge von B, die in den extremen Fällen 0 oder B sein kann, wenn nämlich A mit B oder 0 identisch war, und ihr Komplement ist wieder die ursprüngliche Menge A:

$$B - (B - A) = A.$$

Bei wiederholter Differenzbildung wird stets der Subtrahend als

Teilmenge des Minuenden angenommen. Z. B. bedeutet

$$C-(B-A)$$

die Menge der Elemente von C, die nicht zu $B-A$ gehören, wobei also

$$A \leqq B, \quad B-A \leqq C$$

vorausgesetzt ist.

§ 3. Summe und Durchschnitt.

A und B seien zwei beliebige Mengen. Unter ihrer **Summe**

$$S = \mathfrak{S}(A, B)$$

verstehen wir die Menge der Elemente, die mindestens einer der beiden Mengen angehören; unter ihrem **Durchschnitt**

$$D = \mathfrak{D}(A, B)$$

die Menge der Elemente, die beiden Mengen zugleich angehören.[1]

Ist z. B. (Fig. 1) A die Menge der Punkte des Rechtecks *abcd*, d. h. aller im Innern und auf dem Rande dieses Rechtecks liegenden Punkte der Zeichnungsebene, ebenso B die Rechtecksfläche *pqrs*, so ist ihre Summe die **T**-förmige von *agpqhbcda* umgrenzte Fläche, ihr Durchschnitt die Fläche des kleinen Rechtecks *ghrs*.

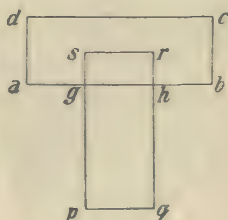


Fig. 1.

Wenn die Mengen A, B kein Element gemein haben, also ihr Durchschnitt die Nullmenge ist ($D = 0$), so nennen wir beide Mengen zueinander **fremd**; in diesem und nur in diesem Falle schreiben wir ihre Summe auch in der Form

$$S = A + B.$$

Wenn $S = A + B$, so ist $A = S - B$, und umgekehrt. Sind A, B endliche Mengen, A aus m, B aus n Elementen bestehend, so besteht $A + B$ aus $m + n$ Elementen.

Im allgemeinen Falle ist D Teilmenge von A und von B, und diese beiden Mengen Teilmengen von S ($D \leqq A \leqq S$, $D \leqq B \leqq S$). Wenn A Teilmenge von B ist, so ist $D = A$, $S = B$, und umgekehrt zieht eine dieser beiden Gleichungen die andere und die Ungleichung $A \leqq B$ nach sich.

Es ist evident, wie man Summen- und Durchschnittsbildung auf mehr als zwei, selbst auf unendlich viele Mengen zu übertragen hat. Beschränken wir uns zunächst auf den Fall eines Systems

---

[1] Die Summe entspricht einem Entweder-oder, der Durchschnitt einem Sowohl-als-auch. In der mathematischen Logik pflegt man Durchschnittsbildung als eine Art Multiplikation aufzufassen.

von endlich vielen Mengen $A_1, A_2, \ldots, A_m$ (die nicht paarweise verschieden zu sein brauchen).

Die Summe

$$S = \mathfrak{S}(A_1, A_2, \ldots, A_m) = \mathfrak{S}_i A_i \qquad (i = 1, 2, \ldots, m)$$

ist die Menge der Elemente, die mindestens einer Menge A_i angehören; der Durchschnitt

$$D = \mathfrak{D}(A_1, A_2, \ldots, A_m) = \mathfrak{D}_i A_i$$

die Menge der Elemente, die allen Mengen A_i zugleich angehören. Nur in dem Falle, daß die Mengen A_i paarweise fremd sind, d. h.

$$\mathfrak{D}(A_i, A_k) = 0 \text{ für } i \neq k,$$

schreiben wir die Summe auch in der Gestalt

$$S = A_1 + A_2 + \ldots + A_m = \sum_i A_i$$

(unter Verwendung des griechischen Summenzeichens Σ statt des deutschen $\mathfrak{S}$).[1]

Jede der beiden Verknüpfungen $\mathfrak{S}$, $\mathfrak{D}$ befolgt das kommutative Gesetz, d. h. die Reihenfolge der Mengen A_i ist gleichgültig. Sie befolgt ferner das assoziative Gesetz, das eine Summe von Summen als Summe, einen Durchschnitt von Durchschnitten als Durchschnitt darzustellen erlaubt und in der Formel

$$\mathfrak{S}[\mathfrak{S}(A_1, A_2, \ldots, A_m), \mathfrak{S}(B_1, B_2, \ldots, B_n), \mathfrak{S}(C_1, C_2, \ldots, C_p), \ldots]$$
$$= \mathfrak{S}(A_1, A_2, \ldots, A_m, B_1, B_2, \ldots, B_n, C_1, C_2, \ldots, C_p, \ldots)$$

sowie einer entsprechenden für Durchschnitte seinen Ausdruck findet. Der Beweis dieser Formeln kann dem Leser überlassen werden. Endlich befolgt jede der beiden Verknüpfungen $\mathfrak{S}$, $\mathfrak{D}$ gegenüber der andern das distributive Gesetz, d. h. eine Summe von Durchschnitten ist als Durchschnitt von Summen darstellbar und umgekehrt. Die Formeln dafür sind:

$$(1) \qquad \mathfrak{D}[\mathfrak{S}(A_1, A_2, \ldots, A_m), \mathfrak{S}(B_1, B_2, \ldots, B_n), \mathfrak{S}(C_1, C_2, \ldots, C_p), \ldots]$$
$$= \mathfrak{S}_{ikl\ldots}[\mathfrak{D}(A_i, B_k, C_l, \ldots)],$$

$$(2) \qquad \mathfrak{S}[\mathfrak{D}(A_1, A_2, \ldots, A_m), \mathfrak{D}(B_1, B_2, \ldots, B_n), \mathfrak{D}(C_1, C_2, \ldots, C_p), \ldots]$$
$$= \mathfrak{D}_{ikl\ldots}[\mathfrak{S}(A_i, B_k, C_l, \ldots)],$$

wobei in den Ausdrücken rechterhand i die Zahlen $1 \ldots m$, k die Zahlen $1 \ldots n$, l die Zahlen $1 \ldots p$ usw. zu durchlaufen hat, die Anzahl der unter dem Summen- resp. Durchschnittszeichen stehenden Mengen also das Produkt der Zahlen $m, n, p, \ldots$ ist.

---

[1] Der Index i unter den Zeichen $\mathfrak{S}$, $\mathfrak{D}$, Σ wird gelegentlich auch wegbleiben können.

Wir geben den Beweis der ersten Formel an, deren linke Seite wir zur Abkürzung mit

$$L = \mathfrak{D}(\mathfrak{S}_A, \mathfrak{S}_B, \mathfrak{S}_C, \ldots)$$

und deren rechte Seite wir mit R bezeichnen. Ein Element von L gehört gleichzeitig zu $\mathfrak{S}_A$, zu $\mathfrak{S}_B$, zu $\mathfrak{S}_C$ usw., also zu mindestens einem A_i, einem B_k, einem C_l usw., also zu mindestens einem $\mathfrak{D}(A_i, B_k, C_l, \ldots)$, also zu R; folglich ist $L \leqq R$. Umgekehrt: ein Element von R gehört zu mindestens einem $\mathfrak{D}(A_i, B_k, C_l, \ldots)$, also gleichzeitig zu A_i, B_k, C_l usw., also gleichzeitig zu $\mathfrak{S}_A$, $\mathfrak{S}_B$, $\mathfrak{S}_C$ usw., also zu L; folglich ist $R \leqq L$ und demnach $L = R$.

Die zweite Formel wäre durch ähnliche Überlegungen zu beweisen, folgt aber auch aus dem in § 4 zu besprechenden Prinzip der Dualität.

Einige Spezialfälle des distributiven Gesetzes mögen noch angegeben werden. So ist

$$\mathfrak{D}[\mathfrak{S}(A_1, A_2, \ldots, A_m), B] = \mathfrak{S}[\mathfrak{D}(A_1, B), \mathfrak{D}(A_2, B), \ldots, \mathfrak{D}(A_m, B)]$$

und

$$\mathfrak{S}[\mathfrak{D}(A_1, A_2, \ldots, A_m), B] = \mathfrak{D}[\mathfrak{S}(A_1, B), \mathfrak{S}(A_2, B), \ldots, \mathfrak{S}(A_m, B)].$$

Sind die Mengen A_i paarweise fremd, so sind es auch die Mengen $\mathfrak{D}(A_i, B)$; in diesem Falle also wird aus der ersten Gleichung

$$\mathfrak{D}(A_1 + A_2 + \ldots + A_m, B) = \mathfrak{D}(A_1, B) + \mathfrak{D}(A_2, B) + \ldots + \mathfrak{D}(A_m, B).$$

Ist $A_1 \leqq B_1$, $A_2 \leqq B_2$, ..., $A_m \leqq B_m$, so ist

$$\underset{i}{\mathfrak{S}} A_i \leqq \underset{i}{\mathfrak{S}} B_i, \quad \underset{i}{\mathfrak{D}} A_i \leqq \underset{i}{\mathfrak{D}} B_i.$$

§ 4. Prinzip der Dualität.

Sind die Mengen $A_1, A_2, \ldots, A_m$ Teilmengen einer Menge M und $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_m$ ihre Komplemente in M, also

$$M = A_i + \bar{A}_i,$$

so ist

$$M = \mathfrak{S}(A_1, A_2, \ldots, A_m) + \mathfrak{D}(\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_m)$$
$$= \mathfrak{D}(A_1, A_2, \ldots, A_m) + \mathfrak{S}(\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_m);$$

denn die Elemente von M gehören entweder mindestens einem A_i oder keinem, d. h. entweder der Summe der A_i oder dem Durchschnitt der $\bar{A}_i$ an, und das gleiche gilt, wenn man die A_i mit den $\bar{A}_i$ vertauscht. Wir können diese Formeln kurz so aussprechen: das Komplement einer Summe ist der Durchschnitt der Komplemente, das Komplement eines Durchschnitts die Summe der Komplemente. Ist also P eine Menge, die aus den Mengen A_i durch wiederholte Summen- und Durchschnittsbildung entsteht, so erhält man ihr Komplement $\bar{P}$, indem man die A_i durch ihre Komplemente $\bar{A}_i$, das Zeichen $\mathfrak{S}$ durch $\mathfrak{D}$ und das Zeichen $\mathfrak{D}$ durch $\mathfrak{S}$ ersetzt. Da ferner

aus $P = Q$, $P < Q$, $P > Q$ resp. $\bar{P} = \bar{Q}$, $\bar{P} > \bar{Q}$, $\bar{P} < \bar{Q}$ folgt, so bleibt jede Gleichung zwischen Mengen richtig, wenn man alle Mengen durch ihre Komplemente ersetzt und die Zeichen $\mathfrak{S}$, $\mathfrak{D}$ vertauscht; jede Ungleichung bleibt richtig, wenn man außerdem noch die Zeichen $<$ $>$ vertauscht. Diesen Sachverhalt, der aus jeder Relation zwischen Mengen eine zweite[1] abzuleiten gestattet, kann man als Dualitätsprinzip bezeichnen, nach dem bekannten Analogon in der Geometrie. Eine identisch, d. h. für beliebige Mengen richtige Relation liefert eine zweite solche auch ohne Übergang zu den Komplementen, also durch bloße Vertauschung der Zeichen $\mathfrak{S}$, $\mathfrak{D}$ und eventuell der beiden Ungleichheitszeichen. Z. B. folgt auf diese Weise die zweite Formel des assoziativen oder distributiven Gesetzes unmittelbar aus der ersten. Als Beispiel für eine Ungleichung zitieren wir die einfache $A \leqq \mathfrak{S}(A, B)$, aus der durch Dualität $A \geqq \mathfrak{D}(A, B)$ folgt.

§ 5. Differenzenketten.

Die schon in § 2 erwähnten, durch wiederholte Differenzbildung entstehenden Mengen bezeichnen wir jetzt unter Weglassung der Klammern in folgender Weise:

$$A_1 - (A_2 - A_3) = A_1 - A_2 + A_3$$
$$A_1 - (A_2 - A_3 + A_4) = A_1 - A_2 + A_3 - A_4$$

usw., definieren also durch Rekursion eine m-gliedrige Differenzenkette vermöge der Formel

$$A_1 - A_2 + A_3 - \cdots \pm A_m = A_1 - (A_2 - A_3 + \cdots \mp A_m).$$

Die Voraussetzung bleibt bestehen, daß der jedesmalige Subtrahend im Minuenden enthalten sein soll. Setzt man

$$A_i - A_{i+1} + \cdots \pm A_m = B_i,$$

so sind die Mengen B_i durch das Gleichungssystem

$$(1)\qquad \left\{\begin{aligned} A_1 &= B_1 + B_2 \\ A_2 &= B_2 + B_3 \\ &\cdots\cdots \\ A_{m-1} &= B_{m-1} + B_m \\ A_m &= B_m \end{aligned}\right.$$

in der Reihenfolge $B_m, B_{m-1}, \ldots, B_2, B_1$ definiert, vorausgesetzt, daß die Auflösung dieses Systems möglich, d. h. daß $B_i \leqq A_{i-1}$ ist. Ist C eine weitere Menge und setzt man

$$\mathfrak{D}(A_i, C) = A_i', \quad \mathfrak{D}(B_i, C) = B_i',$$

---

[1] oder eigentlich unendlich viele, da die Komplemente von der Wahl der Menge M abhängen.

so folgt aus der am Schlusse von § 3 angegebenen speziellen Formel des distributiven Gesetzes, daß das Gleichungssystem (1) auch für die Mengen A_i', B_i' gilt; demnach ist wieder

$$B_1' = A_1' - A_2' + A_3' - \dots \pm A_m'$$

oder

$$\mathfrak{D}(A_1 - A_2 + A_3 - \dots \pm A_m, C) = \mathfrak{D}(A_1, C) - \mathfrak{D}(A_2, C) + \mathfrak{D}(A_3\, C) - \dots \pm \mathfrak{D}(A_m, C),$$

das distributive Gesetz für Differenzenketten.

Gewisse spezielle Differenzenketten befolgen auch das assoziative Gesetz. Ist nämlich

(2) $$A_1 \geqq A_2 \geqq A_3 \geqq \dots \geqq A_m,$$

so ist das System (1) sicher auflösbar, da ja $B_i \leqq A_i \leqq A_{i-1}$. In diesem Fall ist auch das System

$$\begin{aligned} A_1 &= C_1 + C_2 \\ A_2 &= C_2 + C_3 \\ &\dots\dots \\ A_{i-2} &= C_{i-2} + C_{i-1} \\ A_{i-1} &= C_{i-1} + C_i \\ A_i &= C_i \end{aligned}$$

auflösbar und definiert die i-gliedrige Differenzenkette

$$C_1 = A_1 - A_2 + A_3 - \dots \pm A_i.$$

Durch Vergleichung mit (1) findet man der Reihe nach ($i < m$ vorausgesetzt):

$$\begin{gathered} C_i = B_i + B_{i+1}, \\ C_{i-1} + C_i = B_{i-1} + B_i \quad \text{oder} \\ C_{i-1} + B_i + B_{i+1} = B_{i-1} + B_i \quad \text{oder} \\ C_{i-1} + B_{i+1} = B_{i-1}, \\ C_{i-2} + C_{i-1} = B_{i-2} + B_{i-1} \quad \text{oder} \\ C_{i-2} + C_{i-1} = B_{i-2} + C_{i-1} + B_{i+1} \quad \text{oder} \\ C_{i-2} = B_{i-2} + B_{i+1} \quad \text{usw.,} \end{gathered}$$

also

$$B_{i+1} = C_i - B_i = B_{i-1} - C_{i-1} = C_{i-2} - B_{i-2} = \dots$$

Schließlich folgt für gerades i

$$B_{i+1} = B_1 - C_1, \quad B_1 = C_1 + B_{i+1},$$

für ungerades i

$$B_{i+1} = C_1 - B_1, \quad B_1 = C_1 - B_{i+1},$$

also in jedem Falle

$$A_1 - A_2 + A_3 - \dots \pm A_m = \\ = (A_1 - A_2 + \dots \pm A_i) \mp (A_{i+1} - A_{i+2} + \dots \pm A_m).$$

Es ist also die Zerlegung der Glieder der Differenzenkette in zwei Gruppen erlaubt, falls die A_i der Bedingung (2) genügen. Durch

wiederholte Anwendung findet man

$$A_1 - A_2 + \cdots \pm A_m = (A_1 - A_2) + (A_3 - A_4) + \cdots + (A_{m-1} - A_m)$$
$$\text{oder } = (A_1 - A_2) + (A_3 - A_4) + \cdots + A_m,$$

je nachdem m gerade oder ungerade ist; wobei auch das assoziative Gesetz für Summen zur Anwendung gelangt.

Eine einfache Anwendung der Differenzenketten zeigt, daß man Durchschnittsbildung auf Summen- und Differenzenbildung zurückführen kann. Denn ist

$$S = \mathfrak{S}(A, B), \ D = \mathfrak{D}(A, B)$$

und setzt man

$$A' = A - D, \quad B' = B - D,$$

so ist A' die Menge der Elemente, die zu A und nicht zu B gehören, und analog B'. Dann ist aber offenbar

$$S = A' + B' + D = A + B' = A' + B,$$
$$D = A - A' = A - (S - B)$$

oder

$$D = A - S + B = B - S + A.$$

Dieselben Formeln geben die Bestimmung derjenigen Menge A, die mit einer bekannten Menge B einen bekannten Durchschnitt und eine bekannte Summe hat, nämlich

$$A = S - B + D;$$

hier ist die rechte Seite eine spezielle Differenzenkette mit assoziativem Gesetz.

§ 6. Symmetrische Mengen.

Es seien $X_1, X_2, \ldots, X_m$ m beliebige Mengen. Wir kennen bereits Mengen, nämlich Summe und Durchschnitt der X_i, die in kommutativer Weise von den X_i abhängen, d. h. so, daß sie bei einer beliebigen Änderung der Reihenfolge dieser Mengen ungeändert bleiben. Solche Mengen können nach dem Vorbild der Algebra als symmetrische Funktionen von $X_1, X_2, \ldots, X_m$ bezeichnet werden, und genau wie dort spielen gewisse einfachste unter ihnen eine besondere Rolle, die wir symmetrische Grundmengen nennen und folgendermaßen definieren. Es sei (für $i = 1, 2, \ldots, m$) U_i die Menge derjenigen Elemente, die in genau i Mengen des Systems $X_1, X_2, \ldots, X_m$ vorkommen, und

$$A_i = U_i + U_{i+1} + \cdots + U_m$$

die Menge derjenigen Elemente, die in mindestens i Mengen X vorkommen; diese Mengen $A_1, A_2, \ldots, A_m$ sollen die symmetrischen Grundmengen der X heißen. Offenbar ist $A_1 \geqq A_2 \geqq \ldots \geqq A_m$, und A_1 ist die Summe, A_m der Durchschnitt der Mengen X.

Die wesentliche Eigenschaft der symmetrischen Grundmengen st, daß sie aus den X durch bloße Summen- und Durchschnitts-ildung entstehen.[1] Bedeuten nämlich $i, k, l, \ldots$ verschiedene ahlen aus der Reihe $1, 2, \ldots, m$ und setzt man

$$D_i = X_i,\ D_{ik} = \mathfrak{D}(X_i, X_k),\ D_{ikl} = \mathfrak{D}(X_i, X_k, X_l), \ldots, \tag{1}$$

o folgt unmittelbar aus der Definition die Darstellung der symmetri-chen Grundmengen als Durchschnittssummen

$$\left\{\begin{array}{l} A_1 = \mathfrak{S}_{i} D_i \ \ = \mathfrak{S}(D_1, \ D_2, \ D_3, \ \ldots), \\ A_2 = \mathfrak{S}_{ik} D_{ik} = \mathfrak{S}(D_{12}, D_{13}, D_{23}, \ldots), \\ A_3 = \mathfrak{S}_{ikl} D_{ikl} = \mathfrak{S}(D_{123}, D_{124}, \ldots), \\ \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \\ A_m = D_{12\ldots m}. \end{array}\right. \tag{2}$$

Eine zweite Darstellung ergibt sich vermöge des Prinzips der)ualität. Es sei M eine alle X_i enthaltende Menge, $\bar{X}_i = M - X_i$ das omplement von X_i in M, ferner U_0 die Menge der Punkte von M, ie keinem X_i angehören, also

$$M = U_0 + U_1 + \ldots + U_m;$$

enn für $i = 1, 2, \ldots, m$ ist U_i auch die Menge derjenigen Elemente on M, die in genau i Mengen X vorkommen.

Dann ist U_{m-i} die Menge der Elemente von M, die in genau $n - i$ Mengen X, also in genau i Komplementen $\bar{X}$ enthalten sind, und

$$U_0 + U_1 + \ldots + U_{m-i} = \bar{A}_{m-i+1} \quad (i = 1, 2, \ldots, m)$$

ie Menge der Elemente, die in mindestens i Mengen $\bar{X}$ enthalten ind, also sind $\bar{A}_m, \bar{A}_{m-1}, \ldots, \bar{A}_1$ in dieser Reihenfolge die symmetri-chen Grundmengen der Komplemente. Wendet man auf sie die)arstellung (2) an und nimmt davon die Komplemente, so erhält nan für die A_i folgende Darstellung als Summendurchschnitte:

$$S_i = X_i,\ S_{ik} = \mathfrak{S}(X_i, X_k),\ S_{ikl} = \mathfrak{S}(X_i, X_k, X_l), \ldots \tag{3}$$

$$\left\{\begin{array}{l} A_m \ \ \ = \mathfrak{D}_{i} S_i \ \ = \mathfrak{D}(S_1, \ S_2, \ S_3, \ \ldots), \\ A_{m-1} = \mathfrak{D}_{ik} S_{ik} = \mathfrak{D}(S_{12}, S_{13}, S_{23}, \ldots), \\ A_{m-2} = \mathfrak{D}_{ikl} S_{ikl} = \mathfrak{D}(S_{123}, S_{124}, \ldots), \\ \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \\ A_1 \ \ \ = S_{12\ldots m} \end{array}\right. \tag{4}$$

Ist $X_1, X_2, \ldots, X_m$ ein System von Mengen mit den symmetri-chen Grundmengen $A_1, A_2, \ldots, A_m$, ferner $Y_1, Y_2, \ldots, Y_n$ ein zweites

[1] Dagegen erfordern die U_i Differenzbildung.

mit den symmetrischen Grundmengen $B_1, B_2, \ldots, B_n$, so können wir beide Systeme zu einem System

$$Z_1, Z_2, \ldots, Z_{m+n} = X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n$$

vereinigen. Die symmetrischen Grundmengen $C_1, C_2, \ldots, C_{m+n}$ dieses Systems lassen sich durch Summen- und Durchschnittsbildung aus den A, B darstellen (wiederum der Algebra analog). Zu diesem Zweck empfiehlt es sich, die Reihe der symmetrischen Grundmengen nach vorwärts mit Null ($A_{m+1} = A_{m+2} = \ldots = 0$) und nach rückwärts mit

$$A_0 = B_0 = C_0 = M$$

fortzusetzen, wo M irgend eine die Mengen X, Y umfassende Menge ist. Hiernach besteht die Gleichung ($i = 0, 1, 2, \ldots$)

$$(5)\qquad \begin{cases} C_i = \mathfrak{S}[\mathfrak{D}(A_i, B_0), \mathfrak{D}(A_{i-1}, B_1), \mathfrak{D}(A_{i-2}, B_2), \ldots, \\ \qquad \mathfrak{D}(A_1, B_{i-1}), \mathfrak{D}(A_0, B_i)], \end{cases}$$

deren Richtigkeit durch leichte Überlegung einleuchtet. Ein Element der Menge rechts ist nämlich gewiß in einem $\mathfrak{D}(A_{i-k}, B_k)$, also in mindestens $i-k$ Mengen X und mindestens k Mengen Y, also in mindestens i Mengen Z enthalten und gehört demnach zu C_i. Umgekehrt, ein Element von C_i ist in mindestens i Mengen Z enthalten. Es sei genau in p Mengen X und q Mengen Y enthalten, so ist $p + q \geqq i$. Ist $q \leqq i$, so ist $p \geqq i - q$, und das Element ist in $\mathfrak{D}(A_{i-q}, B_q)$ enthalten; ist $q > i$, so ist es in $\mathfrak{D}(A_0, B_i)$ enthalten, gehört also jedenfalls zu der Menge rechterhand in (5).

Umgekehrt lassen sich die A durch die B, C ausdrücken, aber mit Verwendung von Differenzen. Ist nämlich A_{ik} die Menge der Elemente, die in mindestens i Mengen X und in genau k Mengen Y enthalten sind, so ist

$$A_i = A_{i0} + A_{i1} + A_{i2} + \ldots.$$

Andererseits ist

$$A_{ik} = \mathfrak{D}(C_{i+k}, \; B_k - B_{k+1});$$

denn die rechte Seite stellt die Menge der Elemente dar, die in mindestens $i+k$ Mengen Z und in genau k Mengen Y enthalten sind, und das sind zugleich die Elemente von A_{ik}. Also hat man

$$A_i = \mathfrak{D}(C_i, B_0 - B_1) + \mathfrak{D}(C_{i+1}, B_1 - B_2) + \ldots,$$

oder

$$(6)\qquad \begin{cases} A_i = C_i - \mathfrak{D}(C_i, B_1) + \mathfrak{D}(C_{i+1}, B_1) - \mathfrak{D}(C_{i+1}, B_2) + \cdots \\ \qquad + \mathfrak{D}(C_{i+n-1}, B_{n-1}) - \mathfrak{D}(C_{i+n-1}, B_n) + \mathfrak{D}(C_{i+n}, B_n), \end{cases}$$

wo wir die Klammern weggelassen haben, weil jede der Mengen

echterhand die folgende als Teilmenge enthält, also eine spezielle)ifferenzenkette mit assoziativem Gesetz vorliegt.[1]

Wir wollen zwei Systeme von Mengen $X_1, X_2, \ldots, X_m$ und $U_1, U_2, \ldots, U_p$ mit denselben symmetrischen Grundmengen $A_1, A_2, \ldots$ wobei wir uns der Verabredung für die A mit Indices $> m, p$ zu rinnern haben) für den Augenblick **kongruent** nennen und dies urch das Zeichen

$$(\alpha) \qquad X_1, X_2, \ldots, X_m \equiv U_1, U_2, \ldots, U_p$$

ndeuten. Aus (α) und

$$(\beta) \qquad Y_1, Y_2, \ldots, Y_n \equiv V_1, V_2, \ldots, V_q$$

olgt dann wegen (5)

$$(\gamma) \qquad X_1, \ldots, X_m, Y_1, \ldots, Y_n \equiv U_1, \ldots, U_p, V_1, \ldots, V_q,$$

mgekehrt aber wegen (6) aus (β) und (γ) wieder (α).

Also: kongruente Systeme zu kongruenten Systemen hinzugefügt eben kongruente Systeme; kongruente Systeme aus kongruenten ystemen weggelassen geben kongruente Systeme.

Die einfachste Kongruenz ist

$$X + Y \equiv X, Y$$

ür den Fall, daß X und Y fremd sind; denn sowohl das aus der inen Menge $X + Y$ bestehende System wie das aus den beiden lengen X, Y bestehende hat die symmetrischen Grundmengen $X + Y, 0, 0, \ldots$ (bei zwei Mengen besteht das System der symetrischen Grundmengen aus Summe, Durchschnitt und Nullen).)urch wiederholte Anwendung dieser einfachsten Kongruenz ergibt ich, daß aus den Gleichungen

$$(7) \qquad Z_1 = X_1 + Y_1, \quad Z_2 = X_2 + Y_2, \ldots, \quad Z_n = X_n + Y_n$$

ie Kongruenz

$$Z_1, Z_2, \ldots, Z_n \equiv X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n$$

olgt. Zwischen den symmetrischen Grundmengen A, B, C der X, Y, Z bestehen dann die Gleichungen (5), (6); z. B. drückt sich A_1 urch die symmetrischen Mengen B, C der Y, Z folgendermaßen aus mit Beachtung von $B_i \leqq C_i$):

$$(8) \qquad \begin{cases} \mathfrak{S}(Z_1 - Y_1, Z_2 - Y_2, \ldots, Z_n - Y_n) = \\ \quad = C_1 - \mathfrak{D}(C_1, B_1) + \mathfrak{D}(C_2, B_1) - \mathfrak{D}(C_2, B_2) + \ldots \\ \quad\quad + \mathfrak{D}(C_n, B_{n-1}) - \mathfrak{D}(C_n, B_n) \\ \quad = C_1 - B_1 + \mathfrak{D}(C_2, B_1) - B_2 + \ldots + \mathfrak{D}(C_n, B_{n-1}) - B_n. \end{cases}$$

[1] Ein Spezialfall von (6) ist die Bestimmung einer Menge aus ihrer umme und ihrem Durchschnitt mit einer bekannten Menge (§ 5 am Ende).

Endlich machen wir noch folgende Anwendung: wenn

$$(9)\quad \begin{cases}(Z_1 - Y_1) + (Z_2 - Y_2) + \dots + (Z_n - Y_n) \\ \qquad = (W_1 - V_1) + (W_2 - V_2) + \dots + (W_p - V_p),\end{cases}$$

so ist

$$(10)\qquad Z_1, \dots, Z_n, V_1, \dots, V_p \equiv W_1, \dots, W_p, Y_1, \dots, Y_n.$$

Setzt man nämlich

$$Z_i = X_i + Y_i, \quad W_k = U_k + V_k,$$
$$P = X_1 + X_2 + \dots + X_n = U_1 + U_2 + \dots + U_p,$$

so ist

$$Z_1, \dots, Z_n \equiv X_1, \dots, X_n, Y_1, \dots, Y_n$$
$$\equiv P, Y_1, \dots, Y_n,$$

ebenso

$$W_1, \dots, W_p \equiv P, V_1, \dots, V_p,$$

also durch Vereinigung

$$Z_1, \dots, Z_n, P, V_1, \dots, V_p \equiv W_1, \dots, W_p, P, Y_1, \dots, Y_n$$

und daraus durch Weglassung von P (Schluß von (β), (γ) auf (α)) die Formel (10).

Diese Betrachtungen über symmetrische Mengen werden in der Theorie des Inhalts von Punktmengen eine Rolle spielen.

§ 7. Ringe und Körper.

Wir haben im folgenden Mengen zu betrachten, deren Elemente selbst wieder Mengen sind, oder wie wir der Deutlichkeit wegen sagen wollen, Systeme von Mengen. Wir bezeichnen diese Systeme mit großen deutschen Buchstaben, die zu ihnen gehörigen Mengen wie bisher mit großen lateinischen; dabei sehen wir von den trivialen Fällen ab, daß das System gar keine Menge oder nur die Nullmenge enthält, schließen also die Systeme 0 und $\{0\}$ aus.

Ein System von Mengen soll ein Ring[1] heißen, wenn Summe und Durchschnitt je zweier Mengen des Systems wieder dem System angehört. Es gehört dann auch Summe und Durchschnitt endlich vieler Mengen des Systems wieder dem System an.

Wir sind jetzt noch nicht in der Lage, besonders interessante Beispiele geben zu können: als nächstliegendes genüge das System aller Teilmengen einer gegebenen Menge M. Weitere findet der Leser in der Theorie der Punktmengen, z. B. das System aller beschränkten Punktmengen, aller abgeschlossenen Mengen, aller Gebiete.

---

[1] Die Ausdrücke Ring und Körper sind der Theorie der algebraischen Zahlen entnommen, auf Grund einer ungefähren Analogie, an die man nicht zu weitgehende Ansprüche stellen möge.

Ist $\mathfrak{M}$ ein beliebiges System von Mengen, so gibt es sicherlich Ringe über $\mathfrak{M}$ (d. h. die $\mathfrak{M}$ als Teilsystem enthalten) und unter diesen einen kleinsten, der nämlich in jedem andern als Teilsystem enthalten ist. Um dies einzusehen, betrachte man, unter evidenter Ausdehnung der Summen- und Durchschnittsbildung auf Systeme von beliebig vielen Mengen, die Summe N aller Mengen von $\mathfrak{M}$; das System $\mathfrak{N}$ aller Teilmengen von N ist dann ein Ring über $\mathfrak{M}$. Nun ist, wie aus der Definition folgt, der Durchschnitt beliebig vieler Ringe wieder ein Ring. Der Durchschnitt aller Ringe zwischen[1] $\mathfrak{M}$ und $\mathfrak{N}$ (d. h. aller Ringe $\geqq \mathfrak{M}$ und $\leqq \mathfrak{N}$) ist also wieder ein Ring über $\mathfrak{M}$, den wir mit $\mathfrak{R} = \mathfrak{M}_\varrho$ bezeichnen wollen. Er ist in jedem Ringe $\mathfrak{R}'$ über $\mathfrak{M}$ als Teilsystem enthalten; denn der Durchschnitt $\mathfrak{D}(\mathfrak{R}', \mathfrak{N})$ ist ein Ring zwischen $\mathfrak{M}$ und $\mathfrak{N}$, also $\mathfrak{R} \leqq \mathfrak{D}(\mathfrak{R}', \mathfrak{N}) \leqq \mathfrak{R}'$. $\mathfrak{R}$ ist also im oben erklärten Sinne der kleinste Ring über dem System $\mathfrak{M}$.

Dieser kleinste Ring $\mathfrak{R}$ wird erhalten, indem man alle Durchschnittssummen

$$\mathfrak{S}(D_1, D_2, \ldots, D_n)$$

bildet, wo D_i den Durchschnitt aus endlich vielen Mengen des Systems $\mathfrak{M}$ bedeutet. Diese Mengen müssen offenbar jedem Ringe über $\mathfrak{M}$, also auch dem kleinsten Ring angehören; andererseits bilden sie aber selbst schon einen Ring, da nach dem assoziativen Gesetz die Summe, nach dem distributiven der Durchschnitt zweier solcher Mengen wieder von dieser Form ist.

Derselbe Ring wird auch von den Summendurchschnitten

$$\mathfrak{D}(S_1, S_2, \ldots, S_n)$$

gebildet, wo S_i die Summe endlich vieler Mengen von $\mathfrak{M}$ ist.

Ein System von Mengen heiße ein Körper, wenn Summe und Differenz von zwei Mengen des Systems wieder dem System angehört.

Es gehört dann auch der Durchschnitt je zweier Mengen des Systems dem System an; denn setzt man

$$S = \mathfrak{S}(A, B), \quad D = \mathfrak{D}(A, B),$$

so ist (§ 5)

$$D = B - (S - A).$$

Ein Körper ist also jedenfalls ein Ring, aber nicht umgekehrt.

Wie oben erkennt man, daß es über einem gegebenen System $\mathfrak{M}$ von Mengen einen kleinsten Körper $\mathfrak{K} = \mathfrak{M}_\varkappa$ gibt, der offenbar

---

[1] Die anscheinend einfachere Definition: Durchschnitt aller Ringe über $\mathfrak{M}$, vermeiden wir, weil die Gesamtheit aller dieser Ringe nicht als Menge aufzufassen ist (§ 1).

den kleinsten Ring $\mathfrak{R} = \mathfrak{M}_\varrho$ enthält und mit $\mathfrak{R}_\varkappa = \mathfrak{M}_{\varrho\varkappa}$, dem kleinsten Körper über $\mathfrak{R}$, identisch ist. Um ihn zu bilden, setzen wir also das gegebene System bereits als Ring voraus und fügen ihm, falls sie noch nicht in ihm enthalten ist, die Nullmenge hinzu.

Hier werden wir auf Grund der Betrachtungen über symmetrische Mengen leicht zum Ziele gelangen. Betrachten wir einerseits die Differenzensummen

$$S = \mathfrak{S}(X_1 - Y_1, X_2 - Y_2, \ldots, X_m - Y_m)$$

aus Mengen des Ringes $\mathfrak{R}$, wo also X_i und Y_i ($\leqq X_i$) dem Ringe $\mathfrak{R}$ angehören sollen, andererseits die speziellen Differenzenketten

$$K = A_1 - A_2 + A_3 - \ldots \pm A_n,$$

wo $A_1 \geqq A_2 \geqq \ldots \geqq A_n$ und diese Mengen zu $\mathfrak{R}$ gehören sollen. Dann ist jedes K ein S, denn nach dem assoziativen Gesetz ist

$$K = (A_1 - A_2) + (A_3 - A_4) + \ldots;$$

umgekehrt ist jedes S ein K, denn die Formel (8), § 6 stellt S als eine solche Differenzenkette dar, deren Glieder symmetrische Grundmengen der X, Y oder Durchschnitte aus solchen sind und daher wie die Mengen X, Y selbst dem Ringe $\mathfrak{R}$ angehören. Dabei ist wesentlich, daß die symmetrischen Grundmengen nur Summen- und Durchschnitts-, keine Differenzbildung verlangen (§ 6).

Die Mengen S oder K müssen dem kleinsten Körper $\mathfrak{R}$ angehören; wir zeigen jetzt, daß sie selbst schon einen Körper bilden. Die Summe zweier S ist ein S; das folgt aus dem assoziativen Gesetz. Ferner ist aber die Differenz $K - K_1$ zweier K wieder ein K. Setzen wir nämlich

$$K = A_1 - (A_2 - A_3 + \ldots \mp A_n) = A_1 - K_2,$$
$$K_2 = A_2 - A_3 + \ldots \mp A_n,$$

so ist

$$K - K_1 = (A_1 - K_2) - K_1 = A_1 - (K_1 + K_2) = A_1 - K_3,$$

da wir ja schon wissen, daß die Summe zweier K ein K ist. Hierbei ist

$$K_3 = B_2 - B_3 + \ldots \mp B_p,$$

wo $B_2 \geqq B_3 \geqq \ldots \geqq B_p$ und diese Mengen zu $\mathfrak{R}$ gehören. Außerdem können wir $B_2 \leqq A_1$ annehmen, indem wir andernfalls K_3 durch

$$K_3 = \mathfrak{D}(K_3, A_1) = \mathfrak{D}(B_2, A_1) - \mathfrak{D}(B_3, A_1) + \ldots \mp \mathfrak{D}(B_p, A_1)$$

ersetzen, und erhalten demnach $K - K_1$ wieder als spezielle Differenzenkette

$$K - K_1 = A_1 - B_2 + B_3 - \ldots \pm B_p.$$

Die Mengen S oder K bilden also einen Körper und folglich den kleinsten Körper $\mathfrak{R}_\varkappa$.

§ 8. Folgen.

Ein System von Dingen

(1) $$a_1, a_2, a_3, \ldots, a_n, \ldots,$$

worin also jeder natürlichen Zahl 1, 2, 3, ..., n, ... ein Ding entspricht, heißt eine Folge. Die Elemente a_n der Folge brauchen nicht sämtlich verschieden zu sein. Eine Folge ist also keine Menge; sie bestimmt allerdings eine Menge, nämlich die Menge der verschiedenen in ihr auftretenden Elemente, aber diese Elemente treten in der Folge eventuell mehrfach, vielleicht unendlich oft, und überdies in einer bestimmten Anordnung auf. Eine Folge ist eine Zuordnung von Elementen einer Menge zu den natürlichen Zahlen, eine in der Menge der natürlichen Zahlen definierte Funktion (vgl. Kap. II).

Die Zahl n heißt der Index oder die Nummer des Elements a_n.

Aus einer Folge erhalten wir durch Permutation andere Folgen mit denselben Elementen in derselben Häufigkeit. Ist nämlich

$$m_1, m_2, m_3, \ldots, m_n, \ldots$$

eine Folge natürlicher Zahlen, worin jede natürliche Zahl genau einmal auftritt, so ist

$$a_{m_1}, a_{m_2}, a_{m_3}, \ldots, a_{m_n}, \ldots$$

eine Permutation der Folge (1). Lassen wir z. B. jede ungerade Zahl mit der folgenden geraden ihren Platz tauschen, so entsteht die permutierte Folge

$$a_2, a_1, a_4, a_3, \ldots, a_{2n}, a_{2n-1}, a_{2n+2}, a_{2n+1}, \ldots$$

Ist

$$p_1, p_2, \ldots, p_n, \ldots$$

eine Folge wachsender natürlicher Zahlen ($p_1 < p_2 < \ldots$), so nennen wir die Folge

$$a_{p_1}, a_{p_2}, \ldots, a_{p_n}, \ldots$$

eine Teilfolge der Folge (1). Eine Teilfolge der speziellen Gestalt

$$a_p, a_{p+1}, a_{p+2}, \ldots,$$

worin also die Indices alle natürlichen Zahlen von einer gewissen Zahl p an sind, heiße ein Rest der Folge (1). Z. B. ist

$$a_2, a_4, a_6, \ldots, a_{2n}, \ldots$$

oder

$$a_1, a_4, a_9, \ldots, a_{nn}, \ldots$$

eine Teilfolge,

$$a_7, a_8, a_9, \ldots, a_{n+6}, \ldots$$

ein Rest der Folge (1).

Wenn die Elemente einer Teilfolge von (1) eine gewisse Eigenschaft haben, so sagen wir: unendlich viele a_n haben diese Eigenschaft, oder a_n hat diese Eigenschaft unendlich oft. Wenn die Elemente eines Restes von (1) eine gewisse Eigenschaft haben, so

sagen wir: **fast alle** a_n haben diese Eigenschaft, oder a_n hat **schließlich** diese Eigenschaft.

Z. B. ist in der Folge

$$1,\ 1,\ 2,\ \frac{1}{2},\ 3,\ \frac{1}{3},\ \ldots,\ n,\ \frac{1}{n},\ldots$$

unendlich oft $a_n < \frac{1}{100}$, in der Folge

$$1,\ \frac{1}{2},\ \frac{1}{3},\ \ldots,\ \frac{1}{n},\ldots$$

ist schließlich $a_n < \frac{1}{100}$.

Eine Folge $a_1, a_2, \ldots, a_n, \ldots$ und eine zweite Folge $b_1, b_2, \ldots, b_n, \ldots$ können wir zu einer einzigen Gesamtfolge vereinigen, z. B. so:

$$a_1,\ b_1,\ a_2,\ b_2,\ \ldots,\ a_n,\ b_n,\ldots$$

Dasselbe gilt offenbar, wenn an Stelle der einen Folge ein endliches System tritt. Ebenso können drei, vier und beliebig viele Folgen zu einer einzigen Folge vereinigt werden. Ja sogar eine **Folge von Folgen** kann zu einer einzigen Folge vereinigt werden. Man denke sich nämlich diese Folgen zunächst untereinander geschrieben:

$$\begin{array}{l} a_1,\ a_2\ a_3,\ a_4,\ldots \\ b_1,\ b_2,\ b_3,\ b_4,\ldots \\ c_1,\ c_2,\ c_3,\ c_4,\ldots \\ d_1,\ d_2,\ d_3,\ d_4,\ldots \\ \cdot\ \ \cdot\ \ \cdot\ \ \cdot\ \ \cdot\ \ \cdot \end{array}$$

und lese dieses Tableau **diagonal** vom oberen zum linken Rande:

$$a_1;\quad a_2,\ b_1;\quad a_3,\ b_2,\ c_1;\quad a_4,\ b_3,\ c_2,\ d_1;\ldots$$

(hier sind diese Diagonalen durch Semikolons getrennt), womit man eine Gesamtfolge erhält, in der jedes Element der einzelnen Folgen genau einmal auftritt. Man schreibt eine solche Folge von Folgen oder Doppelfolge häufig mit Doppelindices:

$$\begin{array}{l} a_{11},\ a_{12},\ a_{13},\ a_{14},\cdots \\ a_{21},\ a_{22},\ a_{23},\ a_{24},\cdots \\ a_{31},\ a_{32},\ a_{33},\ a_{34},\cdots \\ a_{41},\ a_{42},\ a_{43},\ a_{44},\cdots \\ \cdot\ \ \cdot\ \ \cdot\ \ \cdot\ \ \cdot\ \ \cdot\ \ \cdot \end{array}$$

und die diagonale Anordnung zu einer einzigen Folge

$$a_{11};\quad a_{12},\ a_{21};\quad a_{13},\ a_{22},\ a_{31};\quad a_{14},\ a_{23},\ a_{32},\ a_{41};\cdots$$

besteht dann darin, daß man die a_{mn} in erster Linie nach der Summe der Indices $m+n$, in zweiter Linie nach dem ersten Index m ordnet; d. h. a_{mn} steht vor a_{pq}, falls $m+n < p+q$ oder falls $m+n = p+q$, $m < p$. Es ist klar, daß man eine Folge von Doppel-

folgen oder eine dreifache Folge, deren Elemente mit drei Indices numeriert sind, und ebenso eine Folge von höherer Vielfachheit in gleicher Weise zu einer einfachen Folge anordnen kann.

§ 9. Folgen von Mengen.

Wenn die Elemente einer Folge

$$X_1, X_2, \ldots, X_i, \ldots \tag{1}$$

Mengen sind, so lassen sich auf eine solche Folge von Mengen die Verknüpfungen übertragen, die wir für ein endliches System von Mengen entwickelt haben. Die Summe

$$\mathfrak{S}_i X_i = \mathfrak{S}(X_1, X_2, \ldots) \qquad (i = 1, 2, \ldots)$$

ist die Menge der Elemente, die in mindestens einem X_i vorkommen, der Durchschnitt

$$\mathfrak{D}_i X_i = \mathfrak{D}(X_1, X_2, \ldots)$$

die Menge der Elemente, die zugleich in allen X_i vorkommen. Wenn die X_i paarweise fremd sind, schreiben wir die Summe auch in der Gestalt

$$\sum_i X_i = X_1 + X_2 + \ldots$$

Das assoziative Gesetz besagt, daß

$$\mathfrak{S}[\mathfrak{S}_i X_i, \mathfrak{S}_i Y_i, \ldots] = \mathfrak{S}_i U_i,$$
$$\mathfrak{D}[\mathfrak{D}_i X_i, \mathfrak{D}_i Y_i, \ldots] = \mathfrak{D}_i U_i,$$

wenn $U_1, U_2, \ldots$ eine Gesamtfolge ist, die aus den Folgen $X_1, X_2, \ldots$; $Y_1, Y_2, \ldots$; ... nach Anleitung des § 8 hergestellt wird; diese Folgen können selbst ein endliches System oder eine Folge bilden, auch können statt ihrer endliche Systeme zugelassen werden. Wenn es sich nur um endliche Systeme in endlicher Anzahl handelt, so ist das Gesamtsystem $U_1, U_2, \ldots$ auch nur endlich, und wir kommen auf den früheren Fall (§ 3) zurück. Von dem distributiven Gesetz begnügen wir uns, einen speziellen Fall anzuführen:

$$\mathfrak{D}[\mathfrak{S}_i X_i, \mathfrak{S}_k Y_k] = \mathfrak{S}_{ik} \mathfrak{D}(X_i, Y_k)$$
$$= \mathfrak{S}[\mathfrak{D}(X_1, Y_1), \mathfrak{D}(X_1, Y_2), \mathfrak{D}(X_2, Y_1), \ldots]$$

nebst der entsprechenden Formel, die durch Vertauschung von $\mathfrak{S}$ mit $\mathfrak{D}$ entsteht. Der Durchschnitt aus zwei Summen von Folgen ist also wieder als eine solche Summe darstellbar, da die Doppelfolge der Durchschnitte $\mathfrak{D}(X_i, Y_k)$ sich zu einer einfachen Folge anordnen läßt. Ebenso verhält es sich mit dem Durchschnitt aus drei oder einer endlichen Anzahl solcher Summen; nicht hingegen, wenn

es sich um eine Folge dieser Summen handelt. In diesem Fall würde das distributive Gesetz die Form annehmen

$$\mathfrak{D}[\underset{i}{\mathfrak{S}} X_i, \underset{k}{\mathfrak{S}} Y_k, \underset{l}{\mathfrak{S}} Z_l, \ldots] = \underset{ikl\ldots}{\mathfrak{S}} \mathfrak{D}(X_i, Y_k, Z_l, \ldots),$$

wobei rechterhand aber die Summe über alle Folgen natürlicher Zahlen $i, k, l, \ldots$ zu erstrecken ist; die Menge dieser Folgen aber läßt sich nicht zu einer einfachen Folge anordnen (Kap. III, § 5). Die zweite Formel des distributiven Gesetzes entspringt auch in diesem Falle aus der ersten durch Dualität (Vertauschung von $\mathfrak{S}$ und $\mathfrak{D}$).

Summe und Durchschnitt genügen außerdem dem kommutativen Gesetz, d. h. bleiben bei Permutation der Folge ungeändert, oder sind symmetrische Funktionen der Mengen X_i. Neben ihnen wären wiederum andere Mengen dieser Art und insbesondere die symmetrischen Grundmengen der Folge zu betrachten. Nehmen wir eine die sämtlichen Mengen X_i enthaltende Menge M zu Hülfe; $\bar{X}_i$ sei das Komplement von X_i in M. Für $p = 0, 1, 2, \ldots$ sei dann U_p die Menge der Elemente von M, die genau in p Mengen X (also in unendlich vielen Komplementen $\bar{X}$) vorkommen; V_p die Menge der Elemente von M, die genau in p Komplementen $\bar{X}$ (also in unendlich vielen Mengen X) vorkommen, und W die Menge der Elemente, die in unendlich vielen Mengen X und in unendlich vielen Komplementen $\bar{X}$ vorkommen. Es ist dann

$$M = U_0 + U_1 + U_2 + \ldots + W + V_0 + V_1 + V_2 + \ldots.$$

Weiter ist

$$A_p = U_p + U_{p+1} + \ldots + W + V_0 + V_1 + V_2 + \ldots$$

die Menge der Elemente von M, die in mindestens p Mengen X vorkommen;

$$A_\infty = W + V_0 + V_1 + V_2 + \ldots$$

die Menge derer, die in unendlich vielen X vorkommen;

$$B_\infty = V_0 + V_1 + V_2 + \ldots$$

die Menge derer, die in fast allen X (§ 8) vorkommen, und endlich

$$B_p = V_0 + V_1 + \ldots + V_{p-1} \quad (B_0 = 0)$$

die Menge derer, die in fast allen X bis auf höchstens $p-1$ Ausnahmen vorkommen. Die Mengen $A_1, A_2, \ldots, A_\infty, B_1, B_2, \ldots, B_\infty$, die von der Wahl der Menge M übrigens unabhängig sind (während $A_0 = M$), wollen wir die symmetrischen Grundmengen der X nennen und in der Reihenfolge

$$A_1, A_2, \ldots, A_\infty, B_\infty, \ldots, B_2, B_1$$

geschrieben denken, in der jede alle folgenden als Teilmengen enthält; A_1 ist die Summe, B_1 der Durchschnitt aller X. Man erkennt sofort, daß

$$\bar{B}_1, \bar{B}_2 \ldots, \bar{B}_\infty, \bar{A}_\infty, \ldots, \bar{A}_2, \bar{A}_1$$

die symmetrischen Grundmengen der Komplemente $\bar{X}$ sind.

Auch hier lassen sich die symmetrischen Grundmengen durch bloße Summen- und Durchschnittsbildung aus den X darstellen. Setzt man wieder für verschiedene natürliche Zahlen $i, k, l, \ldots$

$$D_i = X_i, \; D_{ik} = \mathfrak{D}(X_i, X_k), \; D_{ikl} = \mathfrak{D}(X_i, X_k, X_l), \ldots$$
$$S_i = X_i, \; S_{ik} = \mathfrak{S}(X_i, X_k), \; S_{ikl} = \mathfrak{S}(X_i, X_k, X_l), \ldots$$

so ergibt sich unmittelbar aus der Definition, analog zu § 6, (2)

$$A_1 = \underset{i}{\mathfrak{S}} D_i \; = \mathfrak{S}(D_1, \; D_2, \; D_3, \ldots)$$
$$A_2 = \underset{ik}{\mathfrak{S}} D_{ik} = \mathfrak{S}(D_{12}, D_{13}, D_{23}, \ldots)$$
$$A_3 = \underset{ikl}{\mathfrak{S}} D_{ikl} = \mathfrak{S}(D_{123}, D_{124}, \ldots)$$
$$\cdots\cdots\cdots\cdots$$

Wendet man diese Darstellung auf die $\bar{B}_p$ als symmetrische Grundmengen der $\bar{X}_i$ an und nimmt davon die Komplemente, so folgt

$$B_1 = \underset{i}{\mathfrak{D}} S_i \; = \mathfrak{D}(S_1, \; S_2, \; S_3, \; \ldots)$$
$$B_2 = \underset{ik}{\mathfrak{D}} S_{ik} = \mathfrak{D}(S_{12}, S_{13}, S_{23}, \ldots)$$
$$B_3 = \underset{ikl}{\mathfrak{D}} S_{ikl} = \mathfrak{D}(S_{123}, S_{124}, \ldots)$$
$$\cdots\cdots\cdots\cdots$$

Es bleiben noch die beiden Mengen A_∞, B_∞ darzustellen, die gerade die wichtigsten sind. A_∞ war die Menge der Elemente, die unendlich vielen X angehören, B_∞ die Menge derer, die fast allen X (allen bis auf höchstens endlich viele Ausnahmen) angehören. Nach dem Vorbilde der Analysis (vgl. dazu § 11) erscheint es nicht unzweckmäßig, A_∞ als den oberen Limes (limes superior), B_∞ als den unteren Limes (limes inferior) und im Falle der Gleichheit beide als den Limes der Mengenfolge zu bezeichnen.[1] Wir schreiben

$$A_\infty = \operatorname*{Lim\,sup}_i X_i,$$
$$B_\infty = \operatorname*{Lim\,inf}_i X_i$$

und im Falle der Gleichheit beider Mengen (im allgemeinen ist $A_\infty \geqq B_\infty$)

$$A_\infty = B_\infty = \operatorname*{Lim}_i X_i;$$

im letzteren Falle heiße die Mengenfolge konvergent.[2]

---

[1] Diese Limesbegriffe beruhen nur auf der Zugehörigkeit von Elementen zu Mengen; wir werden später andere kennen lernen, die auf gegenseitiger Lage der Elemente beruhen.

[2] Der Index i unter dem Limeszeichen kann gelegentlich wegbleiben.

Zunächst würde sich A_∞ definitionsgemäß als Summe aller Durchschnitte aus unendlich vielen X_i, B_∞ dualistisch als Durchschnitt aller Summen aus unendlich vielen X_i darstellen; auf diese Formeln wollen wir verzichten, da es sich in beiden um ein System von Mengen handelt, das keine Folge bildet. Andererseits ist B_∞ definitionsgemäß die Summe aller Durchschnitte aus fast allen X_i, A_∞ dualistisch der Durchschnitt aller Summen aus fast allen X_i; diese Formeln bevorzugen wir und können dabei noch die Vereinfachung hinzufügen, nur Restfolgen von (1) in Betracht zu ziehen. Setzt man nämlich

$$\Sigma_i = \mathfrak{S}(X_i, X_{i+1}, X_{i+2}, \ldots),$$
$$\Delta_i = \mathfrak{D}(X_i, X_{i+1}, X_{i+2}, \ldots),$$

so ist

$$A_\infty = \mathfrak{D}_i\, \Sigma_i = \mathfrak{D}(\Sigma_1, \Sigma_2, \Sigma_3, \ldots),$$
$$B_\infty = \mathfrak{S}_i\, \Delta_i = \mathfrak{S}(\Delta_1, \Delta_2, \Delta_3, \ldots).$$

In der Tat: ein Element, das fast allen X_i angehört, gehört für ein gewisses i zu Δ_i, also zu $\mathfrak{S}_i \Delta_i$; und ein Element von $\mathfrak{S}_i \Delta_i$ gehört zu einem Δ_i, also zu fast allen X_i. Damit ist die Formel für B_∞ bewiesen, aus der die andere dualistisch folgt.

Wie aus der Definition folgt, bleiben die beiden Mengen $\operatorname{Lim\,sup} X_i$, $\operatorname{Lim\,inf} X_i$ bei Permutation der X_i, außerdem aber auch dann noch unverändert, wenn man eine endliche Anzahl von Mengen hinzufügt, wegläßt oder durch andere Mengen ersetzt.

Ist ferner $Y_1, Y_2, \ldots$ eine Teilfolge von $X_1, X_2, \ldots$, so ergibt sich

$$\operatorname{Lim\,sup} X_i \geqq \operatorname{Lim\,sup} Y_i \geqq \operatorname{Lim\,inf} Y_i \geqq \operatorname{Lim\,inf} X_i.$$

Denn ein Element, das unendlich vielen Y_i angehört, gehört auch unendlich vielen X_i an, und ein Element, das fast allen X_i angehört, gehört auch fast allen Y_i an. Wenn insbesondere $\operatorname{Lim} X_i$ existiert, also die beiden äußeren Mengen der letzten Formel übereinstimmen, so tun es auch die inneren, also existiert auch $\operatorname{Lim} Y_i$ und zwar $= \operatorname{Lim} X_i$. Jede Teilfolge einer konvergenten Mengenfolge ist wieder konvergent und hat denselben Limes.

Wir geben einige Beispiele von konvergenten Folgen. Wenn die X_i eine aufsteigende Folge bilden, d. h.

$$X_1 \leqq X_2 \leqq \ldots,$$

so ist

$$\Delta_i = X_i, \quad \Sigma_i \leqq \Sigma_{i+1},$$

andererseits allgemein $\Sigma_i \geqq \Sigma_{i+1}$, also $\Sigma_i = \Sigma_{i+1} = \Sigma_1$. Also ist

$$A_\infty = \Sigma_1 = \mathfrak{S}(X_1, X_2, \ldots) = B_\infty,$$
$$\operatorname{Lim} X_i = \mathfrak{S}(X_1, X_2, \ldots).$$

Bilden die X_i eine absteigende Folge ($X_1 \geqq X_2 \geqq \ldots$), so findet man ähnlich

$$\operatorname{Lim} X_i = \mathfrak{D}(X_1, X_2, \ldots).$$

Sind die X_i paarweise fremd, so ist $\operatorname{Lim} X_i = 0$.

§ 10. σ-Systeme und δ-Systeme.

Wir nannten (§ 7) ein System von Mengen einen Ring, wenn die Summe und der Durchschnitt zweier (also auch endlich vieler) Mengen des Systems wieder dem System angehört. Wir stellen jetzt die gleiche Forderung für eine Folge von Mengen des Systems; aber hier ist es zweckmäßig, sie in zwei einzelne Forderungen für Summe und Durchschnitt zu trennen.

Ein System von Mengen heiße ein σ-System, wenn die Summe jeder Folge von Mengen des Systems wieder dem System angehört; ein δ-System, wenn der Durchschnitt jeder Folge von Mengen des Systems wieder dem System angehört.

In einem σ-System gehört auch die Summe zweier oder endlich vieler Mengen des Systems dem System an, da man ja

$$\mathfrak{S}(A, B) = \mathfrak{S}(A, B, A, B, A, B, \ldots)$$

schreiben kann; ebenso in einem δ-System der Durchschnitt.

Ein Ring oder Körper, der zugleich ein σ-System ist, heiße ein σ-Ring oder σ-Körper, und entsprechend.

Wie in § 7 überzeugt man sich, daß es zu einem gegebenen Mengensystem $\mathfrak{M}$ ein kleinstes σ-System resp. δ-System gibt, in dem $\mathfrak{M}$ enthalten ist; wir nennen es $\mathfrak{M}_\sigma$ resp. $\mathfrak{M}_\delta$.

Das System $\mathfrak{M}_\sigma$ wird von den Mengen

$$M_\sigma = \mathfrak{S}_i M_i = \mathfrak{S}(M_1, M_2, \ldots)$$

gebildet, wo die M_i Mengen des Systems $\mathfrak{M}$ sind. In der Tat müssen diese Mengen jedem σ-System über $\mathfrak{M}$ angehören; andererseits bilden sie schon selbst ein σ-System, da nach dem assoziativen Gesetz die Summe einer Folge von Mengen M_σ die Summe einer Doppelfolge von Mengen M, also wieder ein M_σ ist.

Das System $\mathfrak{M}_\delta$ besteht ebenso aus den Durchschnitten

$$M_\delta = \mathfrak{D}_i M_i = \mathfrak{D}(M_1, M_2, \ldots)$$

aus Folgen von Mengen des Systems $\mathfrak{M}$.

Ist $\mathfrak{R}$ ein Ring, so ist $\mathfrak{R}_\sigma$ ebenfalls ein Ring (also ein σ-Ring). Da wir bereits wissen, daß die Summe zweier Mengen aus $\mathfrak{R}_\sigma$ wieder zu $\mathfrak{R}_\sigma$ gehört, so ist nur zu zeigen, daß für den Durchschnitt dasselbe gilt. Sind

$$A = \mathfrak{S}_i A_i, \quad B = \mathfrak{S}_k B_k$$

zwei Mengen aus $\mathfrak{R}_\sigma$, so ist nach dem distributiven Gesetz

$$\mathfrak{D}(A, B) = \mathfrak{S}_{ik}\, \mathfrak{D}(A_i, B_k)$$

wieder eine Menge aus $\mathfrak{R}_\sigma$, da die Durchschnitte $\mathfrak{D}(A_i, B_k)$ dem Ringe $\mathfrak{R}$ angehören und eine Doppelfolge bilden.

Die Mengen von $\mathfrak{R}_\sigma$ können, einfacher als im allgemeinen Falle, als Summen aufsteigender Folgen von Mengen aus $\mathfrak{R}$ dargestellt werden. Denn statt $\mathfrak{S}_i A_i$ kann man $\mathfrak{S}_i B_i$ schreiben, wo

$$B_i = \mathfrak{S}(A_1, A_2, \ldots, A_i)$$

wieder dem Ringe $\mathfrak{R}$ angehört und $B_1 \leqq B_2 \leqq B_3 \leqq \ldots$ ist.

Ebenso ist $\mathfrak{R}_\delta$ ein δ-Ring und seine Mengen sind die Durchschnitte absteigender Folgen von Mengen aus $\mathfrak{R}$.

Dagegen braucht, wenn $\mathfrak{K}$ ein Körper ist, $\mathfrak{K}_\sigma$ und $\mathfrak{K}_\delta$ kein Körper zu sein; in der Theorie der Punktmengen werden uns Beispiele dafür begegnen.

Ein σ-Körper ist zugleich ein δ-Körper, aber nicht umgekehrt.

In der Tat, sei $\mathfrak{K}$ ein σ-Körper; wir haben zu zeigen, daß der Durchschnitt jeder Folge von Mengen B_i des Systems $\mathfrak{K}$ wieder zu $\mathfrak{K}$ gehört. Es ist erlaubt, diese Folge absteigend anzunehmen ($B_1 \geqq B_2 \geqq \ldots$). Nun sei M eine zu $\mathfrak{K}$ gehörige Menge über allen B_i (z. B. B_1 selbst). Setzen wir

$$M = A_i + B_i,$$

so ist (§ 4)

$$M = A + B,$$

$$A = \mathfrak{S}(A_1, A_2, \ldots),\; B = \mathfrak{D}(B_1, B_2, \ldots).$$

Auf Grund der Körpereigenschaft gehören nebst den B_i auch die A_i unserem System an, auf Grund der σ-Eigenschaft auch A und auf Grund der Körpereigenschaft wieder B.

Daß ein δ-Körper nicht notwendig ein σ-Körper zu sein braucht, zeigt das Beispiel aller beschränkten Mengen des Raumes. Der obige Beweis kann in umgekehrter Richtung versagen, weil es zwar zu jeder absteigenden, nicht aber zu jeder aufsteigenden Folge von Mengen eines Systems eine sie alle enthaltende Menge M des Systems geben muß.[1] Wenn aber ein δ-Körper über jeder aufsteigenden Folge von Mengen eine sie alle umfassende Menge, z. B. wenn er eine größte Menge enthält, so ist er auch ein σ-Körper.

---

[1] Die Körpereigenschaft ist dem Prinzip der Dualität gegenüber nicht invariant, da der in ihre Definition eingehende Begriff der Differenz die Summe vor dem Durchschnitt bevorzugt. Die Komplemente der Mengen eines Körpers bilden im allgemeinen keinen Körper.

Wir haben in diesem Paragraphen σ-Systeme und δ-Systeme, in § 7 Ringe (ϱ-Systeme) und Körper ($\varkappa$-Systeme) erklärt; ein System kann nun mehrere dieser Eigenschaften zugleich haben. Wenn ein System z. B. gleichzeitig σ-System und δ-System ist, so möge es ein $(\sigma\delta)$-System oder $(\delta\sigma)$-System genannt werden. Ein $(\varrho\sigma)$-System ist ein σ-Ring, ein $(\varkappa\sigma)$-System ein σ-Körper. Eine einfache explizite Darstellung der Mengen solcher Systeme, wie wir sie für Ringe, Körper, σ-Systeme und δ-Systeme angaben, scheint beim Zusammentreffen zweier von den Eigenschaften $\varkappa$, σ, δ nicht vorhanden zu sein. Betrachten wir z. B. das kleinste $(\sigma\delta)$-System über einem gegebenen System $\mathfrak{M}$. In dieses System, das wir $\mathfrak{M}_{(\sigma\delta)} = \mathfrak{M}_{(\delta\sigma)}$ nennen wollen, sind folgende Mengen aufzunehmen: die Mengen M_σ, Summen aus Folgen von Mengen M (des Systems $\mathfrak{M}$); die Mengen $M_{\sigma\delta}$, Durchschnitte aus Folgen von Mengen M_σ; die Mengen $M_{\sigma\delta\sigma}$, Summen aus Folgen von Mengen $M_{\sigma\delta}$ usw. Diese Mengen bilden der Reihe nach das kleinste σ-System $\mathfrak{M}_\sigma$ über $\mathfrak{M}$, das kleinste δ-System $\mathfrak{M}_{\sigma\delta}$ über $\mathfrak{M}_\sigma$, das kleinste σ-System $\mathfrak{M}_{\sigma\delta\sigma}$ über $\mathfrak{M}_{\sigma\delta}$ usw. Aber der Prozeß ist damit nicht beendigt: das ganze System

$$\mathfrak{N} = \mathfrak{S}(\mathfrak{M}, \mathfrak{M}_\sigma, \mathfrak{M}_{\sigma\delta}, \mathfrak{M}_{\sigma\delta\sigma}, \ldots)$$

ist in dem gesuchten System enthalten, weiter aber wieder $\mathfrak{N}_\sigma$, $\mathfrak{N}_{\sigma\delta}$, $\mathfrak{N}_{\sigma\delta\sigma}, \ldots$, die Summe $\mathfrak{P}$ aller dieser Systeme, sodann $\mathfrak{P}_\sigma$, $\mathfrak{P}_{\sigma\delta}$, $\mathfrak{P}_{\sigma\delta\sigma}, \ldots$, deren Summe $\mathfrak{Q}$, schließlich die Summe

$$\mathfrak{S}(\mathfrak{M}, \mathfrak{N}, \mathfrak{P}, \mathfrak{Q}, \ldots),$$

darüber abermals das kleinste σ-System usw. Wir erhalten damit nach einem Bildungsgesetz, das wir später als das der wohlgeordneten Mengen kennen lernen werden, eine Menge von Systemen, die im allgemeinen keine Folge bilden, und deren Gesamtsumme erst das uns vorschwebende System $\mathfrak{M}_{(\sigma\delta)}$ ergeben würde.

Wie aus § 9 hervorgeht, ist der obere Limes einer Folge von Mengen M eine Menge $M_{\sigma\delta}$, der untere eine Menge $M_{\delta\sigma}$.

§ 11. Folgen reeller Zahlen und Funktionen.

Jede reelle Zahl x bestimmt die beiden Mengen:[1]

$$X = \text{Menge der reellen Zahlen} < x,$$

$$\mathfrak{X} = \text{Menge der reellen Zahlen} \leqq x.$$

Wenn die Menge W der reellen Zahlen in zwei (nichtverschwindende) Teilmengen derart zerlegt ist ($W = U + V$), daß jede Zahl von U kleiner als jede Zahl von V ist, so hat bekanntlich entweder die Menge U eine größte oder die Menge V eine kleinste Zahl x; U ist

[1] Die deutschen Buchstaben bedeuten im folgenden Mengen, nicht Mengensysteme.

also entweder ein $\mathfrak{X}$ oder ein X. Diese Eigenschaft, eine Folge der Einführung der irrationalen Zahlen, wird als Stetigkeit (im Dedekindschen Sinne) der Menge W bezeichnet; wir werden in der Theorie der geordneten Mengen darauf zurückzukommen haben.

Die Summe beliebig vieler U ist ein U oder die ganze Zahlenmenge W, der Durchschnitt beliebig vieler U ist ein U oder Null. Die Summe beliebig vieler X ist ein X oder W, der Durchschnitt endlich vieler X ist ein X; der Durchschnitt beliebig vieler $\mathfrak{X}$ ist ein $\mathfrak{X}$ oder Null, die Summe endlich vieler $\mathfrak{X}$ ist ein $\mathfrak{X}$.

Von den Zahlen $x_1, x_2, \ldots, x_m$ sei $y = \min x_i$ die kleinste, $z = \max x_i$ die größte. Für die zugehörigen Mengen gilt dann offenbar

$$Y = \mathfrak{D} X_i, \quad \mathfrak{Y} = \mathfrak{D} \mathfrak{X}_i,$$
$$Z = \mathfrak{S} X_i, \quad \mathfrak{Z} = \mathfrak{S} \mathfrak{X}_i.$$

Betrachten wir sodann eine Zahlenfolge $x_1, x_2, \ldots, x_i, \ldots$ Wenn sie nach unten beschränkt ist, d. h., wenn es Zahlen u gibt, die kein x_i übertreffen ($u \leqq x_i$ für jedes i), so verschwindet $\mathfrak{D} \mathfrak{X}_i$ nicht und es gibt eine Zahl y_1, die größte dieser Zahlen u, für welche

$$\mathfrak{Y}_1 = \mathfrak{D} \mathfrak{X}_i.$$

Diese Zahl y_1, die zu einer Zahlenfolge in derselben Beziehung steht wie das Minimum zu endlich vielen Zahlen, heißt die untere Schranke[1] der Zahlenfolge $x_1, x_2, \ldots$ Überdies ist

$$Y_1 \leqq \mathfrak{D} X_i \leqq \mathfrak{D} \mathfrak{X}_i = \mathfrak{Y}_1,$$

der Durchschnitt der Mengen X_i kann Y_1 oder $\mathfrak{Y}_1$ sein.

Entsprechend ist für eine nach oben beschränkte Zahlenfolge die obere Schranke z_1 durch

$$Z_1 = \mathfrak{S} X_i \leqq \mathfrak{S} \mathfrak{X}_i \leqq \mathfrak{Z}_1$$

definiert.

Für eine nach unten beschränkte Zahlenfolge sei y_i die untere Schranke der Restfolge $x_i, x_{i+1}, x_{i+2}, \ldots$ Wenn die y_i eine nach oben beschränkte Folge bilden, so sei y ihre obere Schranke; y heißt die untere Grenze der Zahlenfolge $x_1, x_2, \ldots$ und man schreibt

$$y = \lim\inf x_i.$$

Auf Grund der Formeln für die Schranken hat man

$$Y_i \leqq \mathfrak{D}(X_i, X_{i+1}, \ldots) \leqq \mathfrak{D}(\mathfrak{X}_i, \mathfrak{X}_{i+1}, \ldots) = \mathfrak{Y}_i,$$
$$Y = \mathfrak{S}\, Y_i \leqq \mathfrak{S}\, \mathfrak{Y}_i \leqq \mathfrak{Y},$$

also

$$Y \leqq \mathfrak{S}_i \mathfrak{D}(X_i, X_{i+1}, \ldots) \leqq \mathfrak{S}_i \mathfrak{D}(\mathfrak{X}_i, \mathfrak{X}_{i+1}, \ldots) \leqq \mathfrak{Y}$$

oder

$$Y \leqq \operatorname{Lim\,inf} X_i \leqq \operatorname{Lim\,inf} \mathfrak{X}_i \leqq \mathfrak{Y}.$$

[1] Die Bezeichnungsweise ist die von M. Pasch; vgl. auch Kap. VII, § 9.

Der untere Limes der Mengenfolge X_i oder $\mathfrak{X}_i$ bestimmt also die Zahl $y = \lim\inf x_i$ in der Weise, daß er mit einer der beiden Mengen Y oder $\mathfrak{Y}$ identisch ist.

In gleicher Weise wird für

$$z = \lim\sup x_i:$$
$$Z \leqq \operatorname{Lim}\sup X_i \leqq \operatorname{Lim}\sup \mathfrak{X}_i \leqq \mathfrak{Z}.$$

Dieser Zusammenhang rechtfertigt auch die Namen, die wir den betreffenden Mengen gegeben haben.

Ist die Zahlenfolge beschränkt (nach unten und nach oben), so existieren beide Zahlen y, z. Wenn beide gleich sind, ist die Zahlenfolge konvergent, $y = z = \lim x_i$; die Folge der Mengen X_i oder $\mathfrak{X}_i$ braucht zwar dann nicht konvergent zu sein, aber ihr oberer und unterer Limes können sich höchstens um das eine Element y unterscheiden. Dagegen folgt aus der Existenz von $\operatorname{Lim} X_i$ oder $\operatorname{Lim} \mathfrak{X}_i$ die Konvergenz der Zahlenfolge.

Betrachten wir nunmehr eine in der beliebigen Menge T definierte reelle Funktion, d. h. jedem Element t von T sei eine reelle Zahl $x = x(t)$ zugeordnet (vgl. Kap. II). Eine solche Funktion bestimmt[1], für jede reelle Zahl u, die beiden Mengen:

$X(u) =$ Menge derjenigen Elemente t, für die $x(t) > u$,

$\mathfrak{X}(u) =$ Menge derjenigen Elemente t, für die $x(t) \geqq u$.

Zur Veranschaulichung möge man etwa T als Menge der reellen Zahlen annehmen und die Funktion $x = x(t)$ in bekannter Weise mittels rechtwinkliger Koordinaten (t Abszisse, x Ordinate) als Punktmenge („Kurve") in der Ebene repräsentieren. Jeder reellen Zahl u entspricht eine der Abszissenachse parallele Gerade $x = u$; die oberhalb dieser Geraden liegenden Punkte der Kurve geben, auf die Abszissenachse projiziert, die Menge $X(u)$ oder $\mathfrak{X}(u)$, je nachdem man die Schnittpunkte der Kurve mit der Geraden ausschließt oder mitzählt.

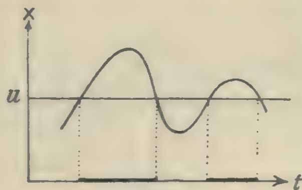


Fig. 2.

Das System der Mengen $X(u)$ und $\mathfrak{X}(u)$, für jedes u, bestimmt seinerseits die Funktion, denn $\mathfrak{X}(u) - X(u)$ ist die Menge derjenigen t, für die $x(t) = u$. Übrigens lassen sich die X durch die $\mathfrak{X}$ ausdrücken

[1] Sie bestimmt natürlich auch, für jedes t, die Menge der reellen Zahlen $< x(t)$ und $\leqq x(t)$; aber deren Betrachtung würde gegenüber der vorigen durchaus nichts Neues ergeben. Statt also, bei festem t, alle u zu sammeln, die der Bedingung $u < x(t)$ oder $u \leqq x(t)$ genügen, sammeln wir bei festem u alle t.

und umgekehrt, so daß schon die Kenntnis des einen von beiden Mengensystemen die Funktion bestimmt.[1] Wenn nämlich k alle natürlichen Zahlen durchläuft, so ist

$$X(u) = \mathfrak{S}_k X\left(u + \frac{1}{k}\right) = \mathfrak{S}_k \mathfrak{X}\left(u + \frac{1}{k}\right),$$
$$\mathfrak{X}(u) = \mathfrak{D}_k X\left(u - \frac{1}{k}\right) = \mathfrak{D}_k \mathfrak{X}\left(u - \frac{1}{k}\right).$$

Jedes $\mathfrak{X}$ ist also Durchschnitt einer Folge von Mengen X oder ein X_δ, jedes X Summe einer Folge von Mengen $\mathfrak{X}$ oder ein $\mathfrak{X}_\sigma$. Die Summe beliebig vieler X (bei festgehaltener Funktion $x(t)$) ist ein X oder die ganze Menge T, der Durchschnitt endlich vieler X ist ein X; der Durchschnitt beliebig vieler $\mathfrak{X}$ ist ein $\mathfrak{X}$ oder Null, die Summe endlich vieler $\mathfrak{X}$ ist ein $\mathfrak{X}$.

Betrachten wir nun mehrere, in derselben Menge T definierte, reelle Funktionen $x_i(t)$ und zwar sogleich eine Folge solcher Funktionen, von der wir voraussetzen, daß die Zahlen $x_i(t)$ für jedes t eine beschränkte Zahlenfolge bilden. Von dieser Zahlenfolge sei $y_1(t)$ die untere, $z_1(t)$ die obere Schranke, ferner

$$y(t) = \lim\inf x_i(t), \quad z(t) = \lim\sup x_i(t).$$

Die diesen Funktionen entsprechenden Mengen seien $Y_1(u)$, $\mathfrak{Y}_1(u)$ usw. Wiederum ist

$$Y_1(u) \leqq \mathfrak{D} X_i(u) \leqq \mathfrak{D}\mathfrak{X}_i(u) = \mathfrak{Y}_1(u),$$
$$Z_1(u) = \mathfrak{S} X_i(u) \leqq \mathfrak{S}\mathfrak{X}_i(u) \leqq \mathfrak{Z}_1(u).$$

Denn wenn $y_1(t) > u$, so ist, für alle i, $x_i(t) > u$ (aber nicht umgekehrt); wenn aber $y_1(t) \geqq u$, so ist, für alle i, $x_i(t) \geqq u$ und umgekehrt. Wenn ferner $z_1(t) > u$, so ist, für mindestens ein i, $x_i(t) > u$ und umgekehrt; ist hingegen, für mindestens ein i, $x_i(t) \geqq u$, so ist auch $z_1(t) \geqq u$ (aber nicht umgekehrt).

Auf Grund dieser Gleichungen gelangt man genau wie zuvor zu den folgenden

$$Y(u) \leqq \operatorname{Lim\,inf} X_i(u) \leqq \operatorname{Lim\,inf} \mathfrak{X}_i(u) \leqq \mathfrak{Y}(u),$$
$$Z(u) \leqq \operatorname{Lim\,sup} X_i(u) \leqq \operatorname{Lim\,sup} \mathfrak{X}_i(u) \leqq \mathfrak{Z}(u).$$

Hiermit sind allerdings die Mengen $Y(u)$ usw. noch nicht bestimmt. Läßt man aber k wieder alle natürlichen Zahlen durchlaufen, so ist

$$Y(u) = \mathfrak{S}_k\, Y\left(u + \frac{1}{k}\right) \leqq \mathfrak{S}_k \operatorname*{Lim\,inf}_i X_i\left(u + \frac{1}{k}\right)$$
$$\leqq \mathfrak{S}_k \operatorname*{Lim\,inf}_i \mathfrak{X}_i\left(u + \frac{1}{k}\right) \leqq \mathfrak{S}_k\, \mathfrak{Y}\left(u + \frac{1}{k}\right) = Y(u),$$

[1] Die $X(u)$ oder $\mathfrak{X}(u)$ können nicht ganz beliebig angenommen werden: wir laden den Leser ein, die notwendigen und hinreichenden Bedingungen aufzustellen.

also

$$Y(u) = \underset{k}{\mathfrak{S}} \underset{i}{\operatorname{Lim\,inf}} X_i\left(u + \frac{1}{k}\right) = \underset{k}{\mathfrak{S}} \underset{i}{\operatorname{Lim\,inf}} \mathfrak{X}_i\left(u + \frac{1}{k}\right),$$

$$\mathfrak{Y}(u) = \underset{k}{\mathfrak{D}} \underset{i}{\operatorname{Lim\,inf}} X_i\left(u - \frac{1}{k}\right) = \underset{k}{\mathfrak{D}} \underset{i}{\operatorname{Lim\,inf}} \mathfrak{X}_i\left(u - \frac{1}{k}\right),$$

$$Z(u) = \underset{k}{\mathfrak{S}} \underset{i}{\operatorname{Lim\,sup}} X_i\left(u + \frac{1}{k}\right) = \underset{k}{\mathfrak{S}} \underset{i}{\operatorname{Lim\,sup}} \mathfrak{X}_i\left(u + \frac{1}{k}\right),$$

$$\mathfrak{Z}(u) = \underset{k}{\mathfrak{D}} \underset{i}{\operatorname{Lim\,sup}} X_i\left(u - \frac{1}{k}\right) = \underset{k}{\mathfrak{D}} \underset{i}{\operatorname{Lim\,sup}} \mathfrak{X}_i\left(u - \frac{1}{k}\right).$$

Um den Aufbau dieser Mengen festzustellen, bezeichnen wir alle $X_i(u)$ mit X, alle $\mathfrak{X}_i(u)$ mit $\mathfrak{X}$ und erinnern uns, daß der obere Limes einer Folge von Mengen M ein $M_{\sigma\delta}$, der untere ein $M_{\delta\sigma}$ ist (§ 10). Danach sind

$$Y(u), \quad \mathfrak{Y}(u), \quad Z(u), \quad \mathfrak{Z}(u)$$

Mengen der Form

$$X_{\delta\sigma}, \quad X_{\delta\sigma\delta}, \quad X_{\sigma\delta\sigma}, \quad X_{\sigma\delta}$$

oder

$$\mathfrak{X}_{\delta\sigma}, \quad \mathfrak{X}_{\delta\sigma\delta}, \quad \mathfrak{X}_{\sigma\delta\sigma}, \quad \mathfrak{X}_{\sigma\delta}.$$

Wir wollen noch für eine Summe zweier Funktionen

$$x_{12}(t) = x_1(t) + x_2(t)$$

die zugehörigen Mengen ermitteln. Die Ungleichung $x_{12}(t) > u$ ist, für jedes einzelne t, mit der Existenz zweier rationaler Zahlen r_1, r_2 gleichwertig, die den Bedingungen

$$x_1(t) > r_1, \quad x_2(t) > r_2, \quad r_1 + r_2 > u$$

genügen (man wähle, wenn $x_1 + x_2 > u$, erst r_1 und dann r_2 den Ungleichungen $x_1 > r_1 > u - x_2$ und $x_2 > r_2 > u - r_1$ entsprechend; wenn umgekehrt zwei solche rationalen Zahlen existieren, so ist $x_1 + x_2 > u$). Demnach ist

$$X_{12}(u) = \mathfrak{S}\,\mathfrak{D}\,[X_1(r_1), X_2(r_2)],$$

die Summe über alle rationalen Zahlen r_1, r_2 erstreckt, deren Summe $> u$; man sieht leicht, daß auch

$$X_{12}(u) = \mathfrak{S}\,\mathfrak{D}\,[\mathfrak{X}_1(r_1), \mathfrak{X}_2(r_2)].$$

Die Menge $\mathfrak{X}_{12}(u)$ erhält man am einfachsten durch Betrachtung ihres Komplements $T - \mathfrak{X}_{12}(u)$, der Menge derjenigen t, wo $x_{12}(t) < u$; diese Ungleichung ist mit der Existenz zweier rationaler Zahlen gleichwertig, für die

$$x_1(t) < \varrho_1, \quad x_2(t) < \varrho_2, \quad \varrho_1 + \varrho_2 < u.$$

Man erhält demgemäß

$$T - \mathfrak{X}_{12}(u) = \mathfrak{S}\,\mathfrak{D}\,[T - \mathfrak{X}_1(\varrho_1), T - \mathfrak{X}_2(\varrho_2)]$$

oder

$$\mathfrak{X}_{12}(u) = \mathfrak{D}\,\mathfrak{S}\,[\mathfrak{X}_1(\varrho_1), \mathfrak{X}_2(\varrho_2)]$$

und ebenso

$$\mathfrak{X}_{12}(u) = \mathfrak{D}\,\mathfrak{S}\,[X_1(\varrho_1), X_2(\varrho_2)],$$

die Durchschnitte über alle rationalen Zahlen ϱ_1, ϱ_2 mit einer Summe $< u$ erstreckt.

Wenn wir für den Augenblick die Summe und den Durchschnitt endlich vieler Mengen M mit M_s, M_d bezeichnen, und wenn wir beachten, daß die genannten rationalen Zahlenpaare in eine Folge angeordnet werden können (Kap. III, § 5), so ist also X_{12} ein $X_{d\sigma}$ oder $\mathfrak{X}_{d\sigma}$, $\mathfrak{X}_{12}$ ein $X_{s\delta}$ oder $\mathfrak{X}_{s\delta}$.

Wir wollen diese Resultate noch etwas vereinfachen. Sagen wir kurz, eine Funktion $x(t)$ ist von der Klasse $(M, \mathfrak{M})$, wenn jedes $X(u)$ ein M und jedes $\mathfrak{X}(u)$ ein $\mathfrak{M}$ ist, wobei M und $\mathfrak{M}$, die natürlich Teilmengen von T bedeuten, gewisse Mengensysteme durchlaufen sollen. Nehmen wir jetzt insbesondere an, daß die M einen σ-Ring und die $\mathfrak{M}$ einen δ-Ring bilden.[1] Suchen wir dann für die bisher abgeleiteten Mengen die einfachste Darstellung: z. B. ist $Y(u)$ ein $M_{\delta\sigma}$ oder $\mathfrak{M}_{\delta\sigma}$, da aber $\mathfrak{M}_\delta$ ein $\mathfrak{M}$ ist, auch ein $\mathfrak{M}_\sigma$; $X_{12}(u)$ ist ein $M_{d\sigma} = M_\sigma = M$. Man erkennt auf diese Weise: die Summe zweier Funktionen der Klasse $(M, \mathfrak{M})$ ist von derselben Klasse; der untere Limes $y(t)$ einer Folge von Funktionen der Klasse $(M, \mathfrak{M})$ ist von der Klasse $(\mathfrak{M}_\sigma, \mathfrak{M}_{\sigma\delta})$, der obere Limes $z(t)$ von der Klasse $(M_{\delta\sigma}, M_\delta)$. Die Summe zweier unterer Limites ist wieder von der Klasse $(\mathfrak{M}_\sigma, \mathfrak{M}_{\sigma\delta})$, da ja auch die $\mathfrak{M}_\sigma$ einen σ-Ring, die $\mathfrak{M}_{\sigma\delta}$ einen δ-Ring bilden. Der Limes $x'(t)$ einer konvergenten Folge von Funktionen $x(t)$ der Klasse $(M, \mathfrak{M})$ ist, da er sowohl ein $y(t)$ wie ein $z(t)$ ist, von der Klasse $(\mathfrak{M}_\sigma, M_\delta)$; der Limes $x''(t)$ einer konvergenten Folge von Funktionen $x'(t)$ ist von der Klasse $(M_{\delta\sigma}, \mathfrak{M}_{\sigma\delta})$ usw.

Nehmen wir noch weiter an, daß die M und $\mathfrak{M}$ Komplemente voneinander seien (jedes $T - M$ ist ein $\mathfrak{M}$, jedes $T - \mathfrak{M}$ ein M). Dann ist $-x(t)$ von der Klasse $(T - \mathfrak{M}, T - M) = (M, \mathfrak{M})$, also von derselben Klasse wie $x(t)$; auch die Differenz zweier solcher Funktionen $x(t)$ ist von derselben Klasse. Der negative untere Limes $-y(t)$ ist von derselben Klasse $(T - \mathfrak{M}_{\sigma\delta}, T - \mathfrak{M}_\sigma) = (M_{\delta\sigma}, M_\delta)$ wie ein oberer Limes; was auch aus der Formel $-y(t) = \limsup\left(-x_i(t)\right)$ hervorgeht. Die Summe zweier Funktionen $z(t)$, oder die Differenz $z(t) - y(t)$ ist von derselben Klasse $(M_{\delta\sigma}, M_\delta)$; insbesondere ist die Menge der Elemente t, wo $z(t) - y(t) > 0$, eine Menge $M_{\delta\sigma}$. Das heißt

---

[1] Beispielsweise, unter Bezugnahme auf spätere Betrachtungen (Kap. IX): die M seien mit den Relativgebieten in T, die $\mathfrak{M}$ mit den in T abgeschlossenen Mengen identisch; die Funktionen der Klasse $(M, \mathfrak{M})$ sind dann die stetigen Funktionen. Hier ist auch die nachherige Annahme erfüllt, daß die M und $\mathfrak{M}$ Komplemente von einander sein sollen.

aber: die Menge der Elemente t, wo eine Folge von Funktionen der Klasse $(M, \mathfrak{M})$ nicht konvergiert, ist ein $M_{\delta\sigma}$, also die Menge derer, wo sie konvergiert, ein $\mathfrak{M}_{\sigma\delta}$.

Wir hatten dabei allerdings vorausgesetzt, daß die $x_i(t)$ für jedes t eine beschränkte Zahlenfolge bilden; diese Einschränkung läßt sich durch eine Abbildung der Menge aller reellen Zahlen auf ein endliches Intervall leicht beseitigen. Ordnen wir z. B. jeder Funktion $x(t)$ die Funktion

$$\xi(t) = \frac{2}{\pi} \operatorname{arctg} x(t) \quad (-1 < \xi(t) < 1)$$

zu. Die Menge der Elemente t, wo $\xi(t) > v$, ist ein $X(u)$ oder T oder 0; ebenso die Menge der t, wo $\xi(t) \geqq v$, ein $\mathfrak{X}(u)$ oder T oder 0. Wenn wir also den Mengen $M, \mathfrak{M}$ nötigenfalls die beiden Mengen T und 0 hinzufügen, so ist mit $x(t)$ auch $\xi(t)$ von der Klasse $(M, \mathfrak{M})$. Ist dann $x_i(t)$ eine beliebige Folge von Funktionen, so braucht $y(t) = \liminf x_i(t)$ nicht für jedes t zu existieren oder es kann, wie man in bekannter Weise zu definieren pflegt, einen der Werte $\pm\infty$ haben, während $\eta(t) = \liminf \xi_i(t)$ immer existiert und entweder gleich $\frac{2}{\pi} \operatorname{arctg} y(t)$ oder aber gleich ± 1 ist. Entsprechendes gilt für $z(t)$ und $\zeta(t)$. Die Konvergenzmenge der $x_i(t)$ ist also die Konvergenzmenge der $\xi_i(t)$ mit Ausschluß derjenigen t, wo $\eta(t) = \zeta(t) = \pm 1$ ist, oder der Durchschnitt der Konvergenzmenge der $\xi_i(t)$ mit der Menge derjenigen t, wo gleichzeitig $\eta(t) > -1$ und $\zeta(t) < 1$. Die Ungleichung $\eta(t) > -1$ ist für eine Menge $\mathfrak{M}_\sigma$ erfüllt, weil $\eta(t)$ wie $y(t)$ von der Klasse $(\mathfrak{M}_\sigma, \mathfrak{M}_{\sigma\delta})$ ist; $\zeta(t) < 1$ für eine Menge $T - M_\delta = \mathfrak{M}_\sigma$; also ist die Konvergenzmenge der Folge $x_i(t)$ eine Menge $\mathfrak{D}(\mathfrak{M}_{\sigma\delta}, \mathfrak{M}_\sigma, \mathfrak{M}_\sigma) = \mathfrak{M}_{\sigma\delta}$.

Wenn die Mengen N ein $(\sigma\delta)$-System bilden und ihre eigenen Komplemente sind[1] ($T - N$ ist wieder ein N), so hat eine Folge von Funktionen der Klasse (N, N) als oberen und unteren Limes wieder eine Funktion der Klasse (N, N); im Konvergenzfalle ist der Limes wieder eine solche Funktion; die Menge der Konvergenzstellen ist eine Menge N.

---

[1] Diese Voraussetzung trifft auf die im Lebesgueschen Sinne meßbaren Mengen zu (Kap. X); die Funktionen der Klasse (N, N) sind die „meßbaren" Funktionen.

Zweites Kapitel.

Mengen und ihre Verknüpfungen: Funktion, Produkt, Potenz.

§ 1. Eindeutige Funktionen.

Aus zwei verschiedenen Elementen a, b können wir die Menge oder das Paar $\{a, b\} = \{b, a\}$ zusammensetzen; beide Elemente treten darin symmetrisch, gleichberechtigt auf. Wir können sie aber auch zu einer unsymmetrischen, das eine Element vor dem anderen bevorzugenden Verbindung zusammenfassen: wir können das geordnete Paar

$$p = (a, b)$$

bilden, das von dem umgekehrt geordneten

$$p^* = (b, a)$$

unterschieden werden soll. Falls die beiden Elemente gleich sind, können wir sie einerseits zur Menge $\{a\}$, andererseits zu dem geordneten Paar (a, a) zusammenfassen, das in diesem Fall mit seiner Umkehrung identisch ist. Zwei geordnete Paare $p = (a, b)$ und $p' = (a', b')$ gelten dann und nur dann als gleich, wenn $a = a'$, $b = b'$.

Die Doppelindices (i, k) an Elementen einer Determinante, die rechtwinkligen Koordinaten (x, y) von Punkten der Ebene sind geordnete Zahlenpaare. Dieser Begriff ist also in der Mathematik fundamental, und die Psychologie würde hinzufügen, daß geordnete, unsymmetrische, selektive Verknüpfung zweier Dinge sogar ursprünglicher ist als ungeordnete, symmetrische, kollektive. Denken, Sprechen, Lesen und Schreiben sind an zeitliche Reihenfolge gebunden, die sich uns aufzwingt, bevor wir von ihr absehen können. Das Wort ist früher da als die Menge seiner Buchstaben, das geordnete Paar (a, b) früher als das Paar $\{a, b\}$.

Übrigens läßt sich, wenn man will, der Begriff des geordneten Paares auf den Mengenbegriff zurückführen. Sind 1, 2 zwei voneinander wie von a und b verschiedene Elemente, so hat das Paar von Paaren

$$\{\{a, 1\}, \{b, 2\}\}$$

genau die formalen Eigenschaften des geordneten Paares (a, b), nämlich Unvertauschbarkeit von a und b im Falle der Verschiedenheit beider Elemente. Diese Unsymmetrie wird hier durch die Verknüpfung mit den Elementen 1, 2 erreicht und kann, da wir nun

einmal sukzessiv und nicht simultan schreiben, durch die Bezeichnung (a, b) ausgedrückt werden, indem wir das mit 1 gepaarte Element an die erste, das mit 2 gepaarte an die zweite Stelle setzen.

Aus zwei nichtverschwindenden Mengen A, B können wir geordnete Paare $p = (a, b)$ bilden, deren erstes Element a ein Element von A, deren zweites Element b ein Element von B ist. Sind beide Mengen endlich und besteht A aus m, B aus n Elementen, so gibt es mn solcher Paare[1]; das legt den Gedanken nahe, in dieser Weise allgemein die Multiplikation von Mengen zu definieren (§ 2).

Zuvor betrachten wir eine Menge P solcher Paare, und zwar von der Beschaffenheit, daß jedes Element a von A in einem und nur einem Paare p von P als erstes Element auftritt. Jedes Element a bestimmt auf diese Weise ein und nur ein Element b, nämlich dasjenige, mit dem es zu einem Paare $p = (a, b)$ verbunden auftritt; dieses durch a bestimmte, von a abhängige, dem a zugeordnete Element bezeichnen wir mit

$$b = f(a)$$

und sagen, daß hiermit in A (d. h. für alle Elemente von A) eine eindeutige Funktion von a definiert sei. Zwei solche Funktionen $f(a)$, $f'(a)$ sehen wir dann und nur dann als gleich an, wenn die zugehörigen Paarmengen P, P' gleich sind, wenn also, für jedes a, $f(a) = f'(a)$ ist.

Umgekehrt ist bei unseren Voraussetzungen ein Element b entweder mit keinem oder einem oder mehreren (möglicherweise unendlich vielen) Elementen a zu einem Paare $p = (a, b)$ verbunden. Ist die Menge P aber so beschaffen, daß auch jedes Element b in genau einem Paare als zweites Element auftritt, so bestimmt auch b ein einziges mit ihm verbundenes Element

$$a = \varphi(b)$$

und wir haben eine zweite, in B definierte eindeutige Funktion. Diese beiden Funktionen heißen zueinander invers oder jede die Umkehrung der anderen, und solche Funktionen werden als umkehrbar eindeutig oder eineindeutig bezeichnet. Zwischen endlichen Mengen ist diese Beziehung offenbar nur möglich, wenn beide Mengen gleichviel Elemente haben, und das legt den Gedanken nahe, auch unendlichen Mengen in diesem Falle gleiche Elementenanzahl zuzuschreiben (Kap. III). Wir nennen zwei Mengen, zwischen denen eine solche umkehrbar eindeutige Beziehung ihrer Elemente möglich ist, äquivalent, in Zeichen:

$$A \sim B \quad \text{oder} \quad B \sim A.$$

[1] Wir lassen das Beiwort „geordnet“ im folgenden wieder weg.

Jede Menge ist mit sich selbst äquivalent, wie aus der Zuordnung jedes Elementes zu sich selbst $(a = f(a))$ hervorgeht. Die Äquivalenz ist eine transitive Beziehung, d. h.

$$\text{aus } A \sim B,\ B \sim C \text{ folgt } A \sim C;$$

denn die umkehrbar eindeutige Beziehung zwischen a und b, b und c vermittelt auch eine solche zwischen a und c.

Indem wir uns eine eingehendere Betrachtung äquivalenter Mengen für Kap. III vorbehalten, lenken wir die Aufmerksamkeit des Lesers vor allem auf die Tatsache, daß eine unendliche Menge A mit einer von ihr verschiedenen Teilmenge B äquivalent sein kann (und sogar immer mit einer solchen äquivalent ist, wie wir sehen werden). In der allerdings etwas provokanten Form ausgesprochen: A hat ebenso viele Elemente wie B, ist dies eine jener „Paradoxien des Unendlichen", an denen der unvorbereitete Intellekt Anstoß nimmt. Z. B. ist die Menge aller natürlichen Zahlen a mit der Menge der geraden natürlichen Zahlen äquivalent, wie die umkehrbar eindeutige Beziehung

$$b = 2a, \quad a = \tfrac{1}{2} b$$

erkennen läßt, die der Menge der Paare

$$(1, 2)\ (2, 4)\ (3, 6) \ldots$$

entspricht. Auch die Menge der durch 1000 teilbaren natürlichen Zahlen $(1000\,a)$, der Quadratzahlen (a^2), der 1000ten Potenzen (a^{1000}), der Potenzen von 1000 (1000^a), der Zahlen a^a, der Primzahlen usw., so spärlich sie auch in der Menge aller natürlichen Zahlen vertreten sein mögen, sind mit dieser Menge äquivalent. Mengen, die mit der Menge der natürlichen Zahlen äquivalent sind, also in die Gestalt einer Folge verschiedener Elemente

$$f(1)\ f(2) \ldots f(a) \ldots$$

gebracht werden können, heißen abzählbar.

Zwei Strecken können stets umkehrbar eindeutig, z. B. projektiv, aufeinander bezogen werden und sind also äquivalent, wobei ihre Längen keine Rolle spielen. Eine Strecke und eine beliebig kleine Teilstrecke, ein Kilometer und ein Millimeter, der Sonnenball und ein Wassertropfen haben in diesem Sinn „gleichviel Punkte".

Die Menge der Punkte $\mathfrak{x}$ einer Geraden ist mit der Menge der reellen Zahlen x, die Menge der Punkte $\mathfrak{p}$ einer Ebene mit der Menge der geordneten Zahlenpaare $p = (x, y)$ äquivalent. Eine eindeutige Funktion $y = f(x)$ gehört zu einer Menge P von Paaren p, die entsprechende Menge $\mathfrak{P}$ der Punkte $\mathfrak{p}$ ist die „Kurve, die durch die Gleichung $y = f(x)$ dargestellt wird".

§ 2. Summe, Durchschnitt, Produkt, Potenz.

Wir ordnen den Elementen i einer (nichtverschwindenden) Menge J vermöge einer eindeutigen Funktion Elemente

$$a_i = f(i)$$

zu. Da sich der Ausdruck Funktion für das Folgende gelegentlich weniger gut eignet, wollen wir diese Zuordnung auch einen Elementenkomplex[1] nennen; die Menge J heiße (in Anlehnung an die Sprache der Funktionentheorie) das Argument des Komplexes oder der Funktion, und ihr einzelnes Element i der Index des zugeordneten Elements a_i.

Wenn wir einzelne Indices andeuten wollen und demgemäß

$$J = \{i, k, l, \ldots\}$$

schreiben, so würde der Komplex unzweideutig durch die Menge der geordneten Paare

$$(1) \qquad \{(i, a_i), (k, a_k), (l, a_l), \ldots\}$$

zu bezeichnen sein; hierbei ist die Reihenfolge gleichgültig. Wenn wir die weniger schleppende Schreibweise

$$(2) \qquad (a_i, a_k, a_l, \ldots)$$

anwenden wollen, so ist zu beachten, daß bei einem individuell gegebenen Komplex, dessen Elemente auch ohne Indices oder mit abweichenden Nummern bezeichnet sein können (etwa $a, b, c, \ldots$ oder $a_1, a_2, a_3, \ldots$), die Zuordnung zu den Indices unersichtlich werden kann; wir verabreden deshalb eine geordnete Schreibweise, wie wir sie bei den geordneten Paaren und stillschweigend schon in Kap. I angewandt haben, indem wir festsetzen, daß die Elemente, die wir angeben, in der Reihenfolge von links nach rechts bestimmten Indices $i, k, l, \ldots$ zugeordnet sein sollen. In dieser geordneten Schreibweise (2) ist also die Reihenfolge nicht mehr gleichgültig; die Komplexe

$$(a, b, c, \ldots) = \{(i, a), (k, b), (l, c), \ldots\}$$

und

$$(b, c, a, \ldots) = \{(i, b), (k, c), (l, a), \ldots\}$$

sind verschieden.[2]

---

[1] Sind die a_i reelle Zahlen, so haben wir einen Zahlenkomplex, den man unter Umständen (nach Einführung geeigneter Rechnungsoperationen) auch eine komplexe Zahl nennt.

[2] Die Theorie der geordneten Mengen (Kap. IV) ist zur Definition einer geordneten Schreibweise noch nicht erforderlich. Wir ordnen nicht die Menge J, sondern nur die wenigen (zwei oder drei!), jedenfalls endlich vielen Elemente von ihr, die wir angeben, oder noch konkreter: wir ordnen eine Kombination von Schriftzeichen (Buchstaben, Klammern, Punkte, Kommata).

Wenn wir den Indices i statt einfacher Elemente **Mengen**

$$A_i = F(i)$$

zuordnen, so sprechen wir von einem **Mengenkomplex**

(3) $$(A_i, A_k, A_l, \ldots).$$

Solche Mengenkomplexe haben wir in Kap. I bereits behandelt für den Fall, daß J die Menge der natürlichen Zahlen $1, 2, \ldots, m$ oder aller natürlichen Zahlen ist.

Wir werden den **Elementenkomplex** (2) in dem **Mengenkomplex** (3) (unter Voraussetzung desselben Arguments J) **enthalten** oder zu ihm **gehörig** nennen, wenn, für jedes i, a_i ein Element von A_i ist. Ein Mengenkomplex enthält dann und nur dann Elementenkomplexe, wenn keine seiner Mengen die Nullmenge ist.

Für einen Mengenkomplex ergibt sich unmittelbar die schon gelegentlich vorweggenommene Erklärung von **Summe** und **Durchschnitt**. Die Summe

$$\overset{J}{\underset{i}{\mathfrak{S}}} A_i = \mathfrak{S}(A_i, A_k, A_l, \ldots)$$

ist die Menge der Elemente, die in mindestens einem A_i (d. h. für mindestens einen Index i in A_i) vorkommen; der Durchschnitt

$$\overset{J}{\underset{i}{\mathfrak{D}}} A_i = \mathfrak{D}(A_i, A_k, A_l, \ldots)$$

die Menge der Elemente, die zugleich in allen A_i vorkommen. Sind die A_i **paarweise fremd**, d. h.

$$\mathfrak{D}(A_i, A_k) = 0 \text{ für } i \neq k,$$

so schreiben wir die Summe auch in der Gestalt

$$\sum_i^J A_i = A_i + A_k + A_l + \cdots$$

Beispiel: i sei eine reelle Zahl $(0 \leqq i \leqq 1)$ und A_i die Menge der reellen Zahlen $\geqq i$. Die Summe ist A_0, der Durchschnitt A_1.

Jede Menge ist die Summe ihrer aus je einem Element bestehenden Teilmengen:

$$J = \sum_i^J \{i\} = \{i\} + \{k\} + \{l\} + \cdots$$

Die Angabe des Arguments über den Zeichen $\mathfrak{S}$, $\mathfrak{D}$ (und nachher $\mathfrak{P}$) gestattet, den Index unter diesen Zeichen durch ein beliebiges Symbol zu ersetzen, z. B.

$$\overset{J}{\underset{i}{\mathfrak{S}}} A_i = \overset{J}{\underset{k}{\mathfrak{S}}} A_k.$$

Indessen kann diese Angabe, wie auch die des Index gelegentlich wegbleiben.

Als das Produkt

$$\mathfrak{P}_i^J A_i = A_i A_k A_l \ldots$$

definieren wir die Menge aller Elementenkomplexe, die in dem Mengenkomplex enthalten sind, d. h. aller Elementenkomplexe $(a_i, a_k, a_l, \ldots)$, für die jedes a_i ein Element von A_i ist; und falls alle Mengen A_i identisch sind, $A_i = A$, so definieren wir das Produkt als die Potenz

$$\mathfrak{P}_i^J A = A^J$$

mit der Basis A und dem Exponenten J. Oder also: das Produkt ist die Menge aller eindeutigen Funktionen $f(i)$, worin, für jedes i, $f(i)$ Element von A_i ist; die Potenz die Menge aller eindeutigen Funktionen $f(i)$, worin, für jedes i, $f(i)$ Element von A ist. Bei der Schreibung des Produkts ohne das Zeichen $\mathfrak{P}$ können auch Multiplikationspunkte verwendet werden.

Ein wichtiges Beispiel für den Potenzbegriff ist das folgende. Die Basis A bestehe aus zwei Elementen 1, 2. Jeder Elementenkomplex oder jede eindeutige Funktion $f(i)$, die nur einen der beiden Werte 1, 2 annehmen darf, zerlegt J in zwei komplementäre Teilmengen

$$J = J_1 + J_2,$$

worin J_1 die Menge derjenigen i ist, für die $f(i) = 1$, und J_2 die Menge derer, für die $f(i) = 2$. Umgekehrt bestimmt jede Teilmenge J_1 von J ihr Komplement J_2 und damit die Funktion $f(i)$. Die Menge der Funktionen $f(i)$ oder die Potenz $\{1, 2\}^J$ ist also mit der Menge aller Teilmengen von J äquivalent.[1]

§ 3. Die Verknüpfungsgesetze.

Zunächst fragen wir, was geschieht, wenn wir das Argument durch eine äquivalente Menge ersetzen. Es sei $J \sim K$, jedem Element k von K entspreche umkehrbar eindeutig ein Element $i = \varphi(k)$ von J, und wir ordnen jetzt dem Element k die Menge $B_k = A_i = A_{\varphi(k)}$ zu, wodurch wir zwei Mengenkomplexe $(A_i, \ldots)$ und $(B_k, \ldots)$ mit den Argumenten J und K erhalten. Eine ganz einfache Überlegung lehrt, daß Summe und Durchschnitt unverändert bleiben und das Produkt in ein äquivalentes übergeht:

---

[1] Eine endliche Menge aus m Elementen hat 2^m Teilmengen (vgl. S. 4).

$$(1) \qquad \begin{cases} \mathfrak{S}_k^K B_k = \mathfrak{S}_i^J A_i, \\ \mathfrak{D}_k^K B_k = \mathfrak{D}_i^J A_i, \\ \mathfrak{P}_k^K B_k \sim \mathfrak{P}_i^J A_i. \end{cases}$$

Vom Produkt ist nicht mehr zu verlangen, da ja Elementenkomplexe mit verschiedenem Argument verschieden sind.

Wir können diese Formeln als die kommutativen Gesetze der drei Verknüpfungen bezeichnen und so aussprechen: zwei Mengenkomplexe mit denselben Mengen in derselben Häufigkeit[1] geben dieselbe Summe, denselben Durchschnitt und äquivalente Produkte. Sie zeigen, daß Summe und Durchschnitt tatsächlich nur von den Mengen des Komplexes, nicht von ihrer Zuordnung zu den Elementen des Arguments abhängen, wobei sogar auch die Häufigkeit keine Rolle spielt; wir können demgemäß auch bei einer Menge von Mengen (wo von Hause aus eine Zuordnung zu einem Argument gar nicht vorliegt) von Summe und Durchschnitt sprechen. Das Produkt hingegen ist von der Zuordnung nicht nur formal, sondern tatsächlich abhängig, und das Zeichen $A B C \ldots$ stellt je nach der verabredeten Zuordnung zu irgendwelchen Indizes im allgemeinen verschiedene, aber äquivalente Mengen dar. Jedes Produkt $A B$ ist mit der Menge der geordneten Paare (a, b) äquivalent, worin a die Menge A und b die Menge B durchläuft.

Auch wenn $K = J$ ist, also eine umkehrbar eindeutige Abbildung von J auf sich selbst, eine Permutation von J vorliegt, sind die bezüglichen Produkte im allgemeinen nur äquivalent, nicht identisch.

Prüfen wir zweitens, wie sich $\mathfrak{S}, \mathfrak{D}, \mathfrak{P}$ bei Ersetzung der Komplexmengen A_i durch äquivalente Mengen B_i (bei ungeändertem Argument) verhalten. Die einfachsten Beispiele zeigen, daß Summe und Durchschnitt hier nicht einmal in äquivalente Mengen überzugehen brauchen[2]; wenn indessen sowohl die A_i als die B_i paarweise fremd sind, so sind die Summen äquivalent:

$$(2) \qquad \sum_i^J A_i \sim \sum_i^J B_i.$$

---

[1] D. h. die Menge der Indices i, für die $A_i = X$, ist äquivalent mit der Menge der Indizes k, für die $B_k = X$, für jede Menge X.

[2] Eine Menge von 4 und eine Menge von 3 Elementen können einen Durchschnitt von 0, 1, 2, 3 Elementen und entsprechend eine Summe von 7, 6, 5, 4 Elementen haben.

Das Produkt geht in jedem Falle in ein äquivalentes über

(3) $$\mathfrak{P}_i^J A_i \sim \mathfrak{P}_i^J B_i.$$

Sind z. B. alle A_i mit derselben Menge A äquivalent, so ist ihr Produkt mit der Potenz A^J äquivalent.

Die assoziativen Gesetze. Das Argument sei jetzt selbst eine Summe paarweise fremder (nichtverschwindender) Summanden:

(4) $$J = \sum_m^M J_m.$$

Bevor wir die hieraus entspringenden assoziativen Gesetze angeben, bemerken wir, daß nach (2) J mit einer Menge P von geordneten Paaren (m, n) äquivalent ist, worin m die Menge M und, bei festem m, n eine mit J_m äquivalente Menge N_m durchläuft. Denn ist P_m die Menge der Paare mit festem m, so ist

$$P = \sum_m^M P_m, \quad P_m \sim N_m \sim J_m,$$

also $P \sim J$. Sind insbesondere alle J_m einer und derselben Menge N äquivalent, so kann man $N_m = N$ setzen und erhält:

(5) $$\sum_m^M J_m \sim MN \quad \text{für} \quad J_m \sim N.$$

Eine Summe paarweise fremder, äquivalenter Summanden ist also einem Produkt äquivalent (wie ein Produkt gleicher Faktoren eine Potenz ist).

Auf Grund der Voraussetzung (4) gelten nun die assoziativen Gesetze

(6) $$\begin{cases} \mathfrak{S}_i^J A_i = \mathfrak{S}_m^M \mathfrak{S}_i^{J_m} A_i, \\ \mathfrak{D}_i^J A_i = \mathfrak{D}_m^M \mathfrak{D}_i^{J_m} A_i, \\ \mathfrak{P}_i^J A_i \sim \mathfrak{P}_m^M \mathfrak{P}_i^{J_m} A_i. \end{cases}$$

Der Zerlegung des Arguments J in Teilargumente J_m entspricht nämlich die Zerlegung des Komplexes $a = (a_i, \ldots)$ in Teilkomplexe b_m, wo b_m nur die den Elementen von J_m zugeordneten a_i zusammenfaßt, und eine Zusammenfassung dieser Teilkomplexe zu einem Gesamtkomplex $b = (b_m, \ldots)$, der umkehrbar eindeutig dem Komplex a entspricht. Z. B. für

$$J = \{1, 2, 3, 4, 5\} = \{1, 2, 3\} + \{4, 5\} = J_1 + J_2$$

ist

$$a = (a_1, a_2, a_3, a_4, a_5),$$

$$b_1 = (a_1, a_2, a_3), \quad b_2 = (a_4, a_5),$$

$$b = (b_1, b_2) = ((a_1, a_2, a_3), \ (a_4, a_5)).$$

Läßt man jedes a_i die Menge A_i durchlaufen (um nur die Produktformel zu beweisen), so durchlaufen a, b_m, b resp. die Mengen

$$\mathfrak{P}_i^J A_i,\ B_m = \mathfrak{P}_i^{J_m} A_i,\ \mathfrak{P}_m^M B_m,$$

und wegen der umkehrbar eindeutigen Beziehung zwischen a und b sind die erste und dritte Menge äquivalent.

Eine für manche Zwecke bequemere Schreibweise erhält man, wenn man J durch eine äquivalente Menge geordneter Paare ersetzt (was ja nach (1) im Sinne der Äquivalenz nichts ändert); wir dürfen annehmen, J sei selbst eine solche Menge und bestehe aus den Paaren (m, n), wo m die Menge M und, bei festem m, n eine Menge N_m durchläuft; $J_m \sim N_m$ ist die Menge der Paare mit festem m. Schreiben wir noch kürzer mn für (m, n) und bezeichnen die diesem Indexpaar oder Doppelindex zugeordnete Menge mit A_{mn}, so ist

$$(7)\quad \begin{cases} \mathfrak{S}_{mn}^J A_{mn} = \mathfrak{S}_m^M \mathfrak{S}_n^{N_m} A_{mn}, \\ \mathfrak{D}_{mn}^J A_{mn} = \mathfrak{D}_m^M \mathfrak{D}_n^{N_m} A_{mn}, \\ \mathfrak{P}_{mn}^J A_{mn} \sim \mathfrak{P}_m^M \mathfrak{P}_n^{N_m} A_{mn}. \end{cases}$$

Spezialfälle des assoziativen Produktgesetzes sind, nebst Formeln wie

$$(AB)C \sim A(BC) \sim ABC,$$

die assoziativen Potenzgesetze. Nimmt man in (6) alle $A_i = A$, so wird

$$(8)\quad A^J \sim \mathfrak{P}_m^M A^{J_m}$$

oder mit Andeutung einiger Elemente von M durch Ziffern

$$A^{J_1 + J_2 + \cdots} \sim A^{J_1} A^{J_2} \ldots .$$

Nimmt man hierin alle J_m äquivalent, $J_m \sim N$, so wird

$$A^J \sim A^{MN},\ \mathfrak{P}_m^M A^{J_m} \sim \mathfrak{P}_m^M A^N = (A^N)^M,$$

$$(9)\quad A^{MN} \sim (A^N)^M.$$

Setzt man in (7) alle $N_m = N$, so wird

$$\mathfrak{P}_{mn}^J A_{mn} \sim \mathfrak{P}_m^M \mathfrak{P}_n^N A_{mn}.$$

Die linke Seite dieser Formel geht nach dem kommutativen Gesetz in eine äquivalente Menge über, wenn man das Argument J, die Menge der Paare $i = (m, n)$, durch das äquivalente Argument J^*,

die Menge der umgekehrten Paare $i^* = (n, m)$ ersetzt und diesem Paare jetzt die Menge A_{mn} zuordnet. Man erhält auf diese Weise

$$\mathfrak{P}_n^N \mathfrak{P}_m^M A_{mn} \sim \mathfrak{P}_m^M \mathfrak{P}_n^N A_{mn}$$

und wenn man alle A_{mn} mit festem m gleich annimmt ($A_{mn} = A_m$):

(10) $$(\mathfrak{P}_m^M A_m)^N \sim \mathfrak{P}_m^M A_m^N$$

oder

$$(A_1 A_2 \ldots)^N \sim A_1^N A_2^N \ldots$$

Die distributiven Gesetze. Während die assoziativen Gesetze sich auf Summen von Summen, Durchschnitte von Durchschnitten, Produkte von Produkten beziehen und diese wieder als Summen, Durchschnitte, Produkte darstellen, also die Form

$$\mathfrak{S}\mathfrak{S} = \mathfrak{S}, \quad \mathfrak{D}\mathfrak{D} = \mathfrak{D}, \quad \mathfrak{P}\mathfrak{P} \sim \mathfrak{P}$$

haben, verknüpfen die distributiven Gesetze je zwei dieser Operationen und gestatten eine Summe von Durchschnitten als Durchschnitt von Summen und umgekehrt, eine Summe von Produkten als Produkt von Summen und umgekehrt, einen Durchschnitt von Produkten als Produkt von Durchschnitten und umgekehrt darzustellen (oder beide wenigstens in Beziehung zu bringen). Wir haben also drei distributive Gesetze, eine Beziehung zwischen

$$\mathfrak{S}\mathfrak{D} \text{ und } \mathfrak{D}\mathfrak{S},$$
$$\mathfrak{S}\mathfrak{P} \text{ und } \mathfrak{P}\mathfrak{S},$$
$$\mathfrak{D}\mathfrak{P} \text{ und } \mathfrak{P}\mathfrak{D},$$

und jedes zerfällt in zwei Gesetze, je nachdem wir z. B. $\mathfrak{S}\mathfrak{D}$ in $\mathfrak{D}\mathfrak{S}$ umformen wollen oder umgekehrt. Das distributive Gesetz zwischen $\mathfrak{S}$ und $\mathfrak{D}$ ist uns in den einfachsten Fällen schon bekannt.

Wir legen die Bezeichnung mit Doppelindices wie in (7) zugrunde. Der Durchschnitt der Mengen

$$S_m = \mathfrak{S}_n^{N_m} A_{mn}$$

besteht aus den Elementen, die für jedes m mindestens einer Menge A_{mn} angehören, ist also die Summe aller Durchschnitte

$$\mathfrak{D}_m^M A_{mn},$$

wenn man darin für n irgend ein Element von N_m setzt. Wir haben also, wenn wir wieder einige Elemente von M durch Ziffern andeuten, einen beliebigen dem Mengenkomplex

$$(N_1, N_2, \ldots, N_m, \ldots)$$

zugehörigen Elementenkomplex

$$q = (n_1, n_2, \ldots, n_m, \ldots)$$

zu nehmen, ihn die Menge

$$Q = \mathfrak{P}_{m}^{M} N_m$$

durchlaufen zu lassen und mit dieser Menge als Argument die Summe über alle Durchschnitte

$$\mathfrak{D}_{m}^{M} A_{m n_m}$$

zu bilden. Wir erhalten damit, und unter Anwendung der Dualität, die beiden distributiven Formeln

$$(11) \qquad \mathfrak{D}_{m}^{M} \mathfrak{S}_{n}^{N_m} A_{mn} = \mathfrak{S}_{q}^{Q} \mathfrak{D}_{m}^{M} A_{m n_m},$$

$$(12) \qquad \mathfrak{S}_{m}^{M} \mathfrak{D}_{n}^{N_m} A_{mn} = \mathfrak{D}_{q}^{Q} \mathfrak{S}_{m}^{M} A_{m n_m}.$$

Genau ebenso[1] beweist man die Formeln

$$(13) \qquad \mathfrak{P}_{m}^{M} \mathfrak{S}_{n}^{N_m} A_{mn} = \mathfrak{S}_{q}^{Q} \mathfrak{P}_{m}^{M} A_{m n_m},$$

$$(14) \qquad \mathfrak{P}_{m}^{M} \mathfrak{D}_{n}^{N_m} A_{mn} = \mathfrak{D}_{q}^{Q} \mathfrak{P}_{m}^{M} A_{m n_m}.$$

Man bemerke, daß hierin rechts nur Produkte mit demselben Argument M wie links auftreten. Wenn wir umgekehrt $\mathfrak{S}\mathfrak{P}$ oder $\mathfrak{D}\mathfrak{P}$ umformen wollen, ist es ratsam, sich auf diesen einfachen Fall, daß alle Produkte dasselbe Argument M haben, zu beschränken. Man findet dann

$$(15) \qquad \mathfrak{S}_{n}^{N} \mathfrak{P}_{m}^{M} A_{mn} \leqq \mathfrak{P}_{m}^{M} \mathfrak{S}_{n}^{N} A_{mn},$$

$$(16) \qquad \mathfrak{D}_{n}^{N} \mathfrak{P}_{m}^{M} A_{mn} = \mathfrak{P}_{m}^{M} \mathfrak{D}_{n}^{N} A_{mn}.$$

Das Ungleichungszeichen in (15) ist nicht vermeidbar; wir illustrieren die Formeln (13)—(16) durch einfache Beispiele, deren Interpretation wir dem Leser überlassen können.

$$\mathfrak{S}(A, B) . \mathfrak{S}(C, D) = \mathfrak{S}(AC, BC, AD, BD),$$
$$\mathfrak{D}(A, B) . \mathfrak{D}(C, D) = \mathfrak{D}(AC, BC, AD, BD),$$
$$\mathfrak{S}(AC, BD) \leqq \mathfrak{S}(A, B) . \mathfrak{S}(C, D),$$
$$\mathfrak{D}(AC, BD) = \mathfrak{D}(A, B) . \mathfrak{D}(C, D).$$

---

[1] Da die distributiven Gesetze zwischen $\mathfrak{P}$ und $\mathfrak{S}$, $\mathfrak{P}$ und $\mathfrak{D}$ kaum eine Anwendung finden, beschränken wir uns auf eine kurze Angabe.

Daß sich Durchschnitt und Summe gegenüber dem Produkt verschieden verhalten, daß man also in einer Formel mit den Zeichen $\mathfrak{D}$, $\mathfrak{S}$, $\mathfrak{P}$ nicht $\mathfrak{D}$ und $\mathfrak{S}$ vertauschen darf, ist keine Abweichung vom Gesetze der Dualität. Denn dieses beruht auf Komplementbildung; aber das Komplement eines Produktes ist nicht das Produkt der Komplemente, d. h. für

$$M = A_i + B_i$$

ist

$$M^J \geqq \mathfrak{P}_i^J A_i + \mathfrak{P}_i^J B_i,$$

im allgemeinen mit Ausschluß des Gleichheitszeichens.

§ 4. Nichteindeutige Funktionen.

Wenn wir aus zwei Mengen A, B eine Menge geordneter Paare $p = (a, b)$ bilden, so erhalten wir eine eindeutige Funktion $f(a)$, falls jedes Element a in genau einem Paare p als erstes Element auftritt. Lassen wir diese Voraussetzung fallen und wählen die Menge P beliebig, so kann ein Element a jetzt in einem Paare p, aber auch in mehreren oder keinem vorkommen; es kann, wie wir uns ausdrückten, mit einem oder mehreren oder keinem Element b verbunden sein. Bezeichnet $b = f(a)$ ein mit a verbundenes Element oder ein Bild von a (nicht das Bild), so kann a ein oder mehrere oder kein Bild $f(a)$ haben; diese Funktion ist also im allgemeinen nicht eindeutig, sie ist für gewisse Elemente a nulldeutig (d. h. überhaupt nicht definiert) und für die übrigen ein- oder mehrdeutig. Das gleiche ist über die inverse Funktion $a = \varphi(b)$ zu sagen, die eins der mit b verbundenen Elemente a oder ein Urbild von b bezeichnen soll. Auch die zu einer eindeutigen Funktion inverse Funktion ist im allgemeinen nicht eindeutig.

Indessen bestimmt, bei gegebener Paarmenge P, jede Teilmenge X von A eindeutig die Menge

$$F(X)$$

aller Bilder aller Elemente von X; wir nennen dies die Bildmenge von X. Ebenso bestimmt jede Teilmenge Y von B eindeutig die Menge

$$\Phi(Y)$$

aller Urbilder aller Elemente von Y, die Urbildmenge von Y. Insbesondere bestimmt ein einzelnes Element a die Menge $F(\{a\})$ seiner Bilder[1] und ein einzelnes Element b die Menge $\Phi(\{b\})$ seiner

[1] Diese Menge einfach mit $F(a)$ zu bezeichnen, könnte in dem Fall zu Kollisionen führen, daß ein Element von A zugleich eine Menge wäre, die ein anderes Element enthielte, z. B. wenn A die Menge der Elemente und Teilmengen einer anderen Menge ist. Man pflegt diesen in praxi allerdings ungewöhnlichen Fall stillschweigend zu ignorieren.

Urbilder. Jede Menge $F(X)$ ist eine Teilmenge von B, kann aber auch Null sein; ebenso ist $0 \leqq \Phi(Y) \leqq A$.

Wir haben nicht vorausgesetzt, daß alle Elemente von A Bilder haben; die Menge derer, die Bilder haben, sei A_0, und ebenso sei für $X \leqq A$

$$X_0 = \mathfrak{D}(X, A_0)$$

die Menge der Elemente von X, die Bilder haben. Es ist also $F(X) = F(X_0)$. Desgleichen sei B_0 und $Y_0 = \mathfrak{D}(Y, B_0)$ die Menge der Elemente von B und Y, die Urbilder haben, also $\Phi(Y) = \Phi(Y_0)$.

Offenbar ist

$$X_0 \leqq \Phi(F(X_0)),$$
$$Y_0 \leqq F(\Phi(Y_0)).$$

Um z. B. die zweite Formel zu beweisen: ein Element b von Y_0 hat ein Urbild a in $\Phi(Y_0)$, also ist b ein Bild von a und gehört zu $F(\Phi(Y_0))$. Umgekehrt läßt sich nichts schließen: ein Element b von $F(\Phi(Y_0))$ hat ein Urbild a in $\Phi(Y_0)$, a hat ein Bild b' in Y_0, aber b kann von b' verschieden sein und braucht nicht zu Y_0 zu gehören. Wenn aber die Funktion $f(a)$ in A_0 eindeutig ist, so ist $b = b'$; dann tritt also in der zweiten Formel das Gleichheitszeichen ein:

$$Y_0 = F(\Phi(Y_0)).$$

Wir können die obigen Formeln in der Gestalt

$$\mathfrak{D}(X, A_0) \leqq \Phi(F(X)),$$
$$\mathfrak{D}(Y, B_0) \leqq F(\Phi(Y))$$

schreiben; in der zweiten tritt das Gleichheitszeichen ein, wenn $f(a)$ in A nicht mehrdeutig (d. h. wenn sie null- oder eindeutig) ist.

Haben wir ein System oder einen Komplex von Mengen $X_1, X_2, \ldots$ (Teilmengen von A), und ist

$$X = \mathfrak{S}(X_1, X_2, \ldots), \quad X' = \mathfrak{D}(X_1, X_2, \ldots)$$

ihre Summe und ihr Durchschnitt, so ist

$$F(X) = \mathfrak{S}(F(X_1),\ F(X_2), \ldots),$$
$$F(X') \leqq \mathfrak{D}(F(X_1),\ F(X_2), \ldots).$$

Ebenso ist für

$$Y = \mathfrak{S}(Y_1, Y_2, \ldots), \quad Y' = \mathfrak{D}(Y_1, Y_2, \ldots)$$
$$\Phi(Y) = \mathfrak{S}(\Phi(Y_1), \Phi(Y_2), \ldots),$$
$$\Phi(Y') \leqq \mathfrak{D}(\Phi(Y_1),\ \Phi(Y_2), \ldots).$$

Wir beweisen wieder die zweite Formelgruppe; die Summengleichung ist evident. Ist a ein Element von $\Phi(Y')$, so hat a ein Bild b in Y', b gehört zu allen Y_i und a zu allen $\Phi(Y_i)$. Umgekehrt läßt sich wiederum nichts schließen: ist a ein Element des rechterhand stehenden Durchschnitts, also ein Element von jedem $\Phi(Y_i)$,

so hat a ein Bild b_i in Y_i, aber diese Bilder können verschieden sein und brauchen nicht zu Y' zu gehören. Ist jedoch wieder die Funktion $f(a)$ in A_0 eindeutig, so sind alle b_i identisch, $b_i = b$, und b gehört zu Y', also a zu $\Phi(Y')$. Es gilt also die Formel

$$\Phi(Y') = \mathfrak{D}(\Phi(Y_1), \Phi(Y_2), \ldots),$$

wenn die Funktion $f(a)$ in A_0 eindeutig oder in A nicht mehrdeutig ist. In diesem Fall ist für $\mathfrak{D}\, Y_i = 0$ also auch $\mathfrak{D}\, \Phi(Y_i) = 0$; sind insbesondere die Y_i paarweise fremd, so sind es auch die $\Phi(Y_i)$ und die Summenformel lautet dann

$$\Phi(Y_1 + Y_2 + \ldots) = \Phi(Y_1) + \Phi(Y_2) + \ldots$$

Drittes Kapitel.

Kardinalzahlen oder Mächtigkeiten.

§ 1. Äquivalenz und Kardinalzahl.

Wenn man eine Menge von Äpfeln mit einer Menge von Birnen in bezug auf die Anzahl der Gegenstände vergleichen will, so geschieht dies auf dem primitiven Standpunkt in der Weise, daß man einen Apfel mit einer Birne zusammenlegt, dann einen zweiten Apfel mit einer zweiten Birne, und dieses Verfahren bis zu seinem Ende fortsetzt, d. h. bis die eine von beiden Mengen erschöpft ist. Ist damit gleichzeitig auch die andere erschöpft, so haben wir ebenso viele Äpfel wie Birnen. Das ist nun gar nichts anderes als die Bildung einer Menge von Paaren $p = (a, b)$, worin jeder Apfel a und jede Birne b in höchstens einem Paar vorkommt; und gelingt es insbesondere die Paarmenge P so zu bilden, daß jedes a und jedes b in genau einem Paare vorkommt, so haben wir gleiche Anzahl von Äpfeln und Birnen oder (Kap. II, § 1) die Äquivalenz beider Mengen A und B konstatiert.

Wenn aber die Äpfel und die Birnen sich an verschiedenen Orten befinden und ein Transport der einen Menge zu der anderen mit Schwierigkeiten verknüpft ist, so wird der erfinderische Menschengeist auf der nächsthöheren Stufe sich einer vermittelnden Menge bequem transportabler Gegenstände, seien es Steine, Muscheln, Holzstücke, bedienen und aus den Äquivalenzen $A \sim C$, $B \sim C$ die Äquivalenz $A \sim B$ erschließen. Endlich aber wird auch dieser Erdenrest noch überwunden und an Stelle der vermittelnden Menge tritt ein System gesprochener, geschriebener oder gedachter Zeichen, der

Zahlzeichen 1, 2, . . . Das Vergleichen wird damit zum Zählen, und äquivalente Mengen erhalten nun eine gemeinsame Eigenschaft, die Anzahl ihrer Elemente.

Diese Bemerkungen, die weder nach der psychologischen noch nach der kulturgeschichtlichen Seite hin irgendwelchen Anspruch erheben wollen, sollen nur verständlich machen, daß die Äquivalenz die natürliche Grundlage der Vergleichung von Mengen ist und daß mit ihrer Hilfe sogar der anscheinend paradoxe Versuch unternommen werden konnte, auch unendliche Mengen zu zählen.

Wir wissen bereits, daß zwei Mengen, die einer dritten äquivalent sind, auch miteinander äquivalent sind, oder daß die Äquivalenz das transitive Gesetz:

$$\text{aus } A \sim B,\ B \sim C \text{ folgt } A \sim C$$

erfüllt. Mengen eines Systems, die einer gegebenen Menge und damit auch untereinander äquivalent sind, haben etwas Gemeinsames, das im Falle endlicher Mengen die Anzahl der Elemente ist und das man auch im allgemeinen Falle die Anzahl oder Kardinalzahl oder Mächtigkeit nennt. Über die absolute Beschaffenheit dieses neu eingeführten Etwas kann man allerhand verschiedene Auffassungen hegen. G. Cantor definiert die Mächtigkeit einer Menge als den Allgemeinbegriff, der durch Abstraktion von der individuellen Beschaffenheit[1] ihrer Elemente entsteht. B. Russell definiert sie geradezu als die Gesamtheit oder Klasse „aller" mit jener Menge äquivalenten Mengen; dies halten wir bei der uferlosen und antinomischen Beschaffenheit dieser Klasse für bedenklich. Wenn wir analoge Beispiele aus anderen Gebieten der Mathematik herbeiziehen, wird die gegenwärtige Situation nicht klarer; denn wenn wir kongruenten Punktpaaren eine gemeinsame „Entfernung", parallelen Geraden eine gemeinsame „Richtung", ähnlichen Figuren eine gemeinsame „Form" beilegen, so können ja diese Begriffe außerdem wirklich durch Strecken, Winkel oder Zahlen präzisiert werden. Andererseits könnte man den Begriff Mächtigkeit freilich entbehren und alles auf die Betrachtung äquivalenter Mengen beschränken, worunter aber die Bequemlichkeit des Ausdrucks erheblich leiden würde. Übrigens ist darauf aufmerksam zu machen, daß die genannten Schwierigkeiten auch schon bei endlichen Mengen bestehen, wo es ja an verschiedenen Auffassungen des Zahlbegriffes

[1] und der Anordnung; von dieser brauchen wir nicht abzusehen, da wir die Ordnung gar nicht als ursprünglichen Bestandteil in den Mengenbegriff aufgenommen haben, sondern umgekehrt aus einer Menge erst durch zusätzliche Bestimmungen eine geordnete Menge hervorgehen lassen. Dies illustriert übrigens unsere Bemerkung (S. 32), daß geordnete Mengen psychologisch früher da sind als ungeordnete.

durchaus nicht mangelt. Wir werden uns einfach auf den formalen Standpunkt stellen und sagen: einem System von Mengen A ordnen wir eindeutig ein System von Dingen $\mathfrak{a}$ zu derart, daß äquivalenten Mengen und nur solchen dasselbe Ding entspricht, d. h. daß aus $A \sim B$ immer $\mathfrak{a} = \mathfrak{b}$ folgt und umgekehrt. Diese neuen Dinge oder Zeichen nennen wir Kardinalzahlen oder Mächtigkeiten; wir sagen: A hat die Mächtigkeit $\mathfrak{a}$, A ist von der Mächtigkeit $\mathfrak{a}$, $\mathfrak{a}$ ist die Mächtigkeit von A, wohl auch (indem wir $\mathfrak{a}$ als Zahlwort verwenden) A hat $\mathfrak{a}$ Elemente.

Ist n eine natürliche Zahl[1], so ordnen wir einer endlichen, aus n Elementen bestehenden Menge als Mächtigkeit die Zahl n zu. Der Nullmenge 0 ordnen wir die Zahl 0 als Mächtigkeit zu.

Der Menge der natürlichen Zahlen 1, 2, 3, ... geben wir die Mächtigkeit $\aleph_0$ (gesprochen Alef-Null). Die Mengen dieser Mächtigkeit, die also mit der Menge der natürlichen Zahlen äquivalent oder als Folgen verschiedener Elemente darstellbar sind, heißen abzählbar, wie bereits erwähnt wurde.

Der Menge der reellen Zahlen oder der damit äquivalenten Menge der Punkte einer geraden Linie geben wir die Mächtigkeit $\aleph$ (Alef), die auch (mit einem später zu rechtfertigenden Ausdruck) die Mächtigkeit des Kontinuums genannt wird.

§ 2. Vergleichung von Kardinalzahlen.

Wir haben die Gleichheit von Kardinalzahlen durch die Äquivalenz zugehöriger Mengen definiert. Wenn wir nun zwei nicht äquivalente Mengen A, B, also zwei verschiedene Mächtigkeiten $\mathfrak{a}$, $\mathfrak{b}$ haben, ist es dann möglich, die eine von ihnen als die größere, die andere als die kleinere zu charakterisieren? Oder, wie man sich ausdrückt, haben die Kardinalzahlen Größencharakter?

Wenn wir eine Menge Äpfel und eine Menge Birnen durch Zusammenlegung zu Paaren vergleichen, so waren ja nur drei Fälle möglich. Entweder werden beide Mengen gleichzeitig erschöpft, oder es bleiben Birnen übrig, oder es bleiben Äpfel übrig. Im ersten Fall ist $A \sim B$, im zweiten ist A mit einer echten (d. h. von B verschiedenen) Teilmenge von B, im dritten B mit einer echten Teilmenge von A äquivalent. Für die endlichen Anzahlen $\mathfrak{a}$, $\mathfrak{b}$ gilt nach

[1] Die vollständige Mengenlehre hat auch die Theorie der endlichen Mengen ab ovo zu entwickeln, also die gewöhnliche Arithmetik und ihre Erweiterungen zu begründen. Wir verzichten bei unserer Darstellung auf diesen Teil des Programms, der ja auch vor und außerhalb der Mengenlehre genügend behandelt worden ist, und setzen das System der natürlichen, ganzen, rationalen, reellen und komplexen Zahlen als bekannt voraus.

der bekannten Bedeutung der Zeichen $<$ (kleiner) und $>$ (größer) im ersten Fall $\mathfrak{a} = \mathfrak{b}$, im zweiten $\mathfrak{a} < \mathfrak{b}$, im dritten $\mathfrak{a} > \mathfrak{b}$.

Wollte man nach diesem Vorbild stets, falls A mit einer echten Teilmenge von B äquivalent ist, $\mathfrak{a} < \mathfrak{b}$ definieren, so würde man etwas höchst Unzweckmäßiges begehen, nämlich zulassen, daß eine Kardinalzahl kleiner als sie selbst sein könnte. Unsere Beispiele (Kap. II, § 1) zeigten ja, daß eine Menge recht wohl mit einer ihrer echten Teilmengen äquivalent sein kann, z. B. die Menge der natürlichen Zahlen n mit der Menge der Quadratzahlen n^2. Diese Eigenschaft kann offenbar nur unendlichen Mengen zukommen und kommt ihnen, wie wir sehen werden (§ 4), auch stets zu. Wenn wir also den Zeichen $= \; < \; >$ die übliche Bedeutung lassen und insbesondere verlangen wollen, daß von den drei Fällen

$$\mathfrak{a} = \mathfrak{b}, \quad \mathfrak{a} < \mathfrak{b}, \quad \mathfrak{a} > \mathfrak{b}$$

immer nur einer eintreten kann, so werden wir darauf verzichten müssen, jeder echten Teilmenge von A eine Kardinalzahl $< \mathfrak{a}$ zu geben; wir müssen das geheiligte Axiom „totum parte majus" verletzen, wie wir uns überhaupt darauf gefaßt machen müssen, daß die Rechnung mit unendlichen Kardinalzahlen in vielen Punkten von der mit endlichen abweichen wird, ohne daß darin der geringste Einwand gegen diese unendlichen Zahlen zu erblicken wäre.

Wir sind durch die Vergleichung von Mengen darauf geführt worden, neben einer Menge zugleich ihre Teilmengen in Betracht zu ziehen. Bezeichnet A' irgend eine Teilmenge ($0 \leqq A' \leqq A$) von A, B' eine von B, so sind zwischen den beiden Mengen offenbar vier Beziehungen möglich, von denen eine und nur eine eintreten muß, die also eine „vollständige Disjunktion" bilden:

(1) Es gibt ein $A' \sim B$, und es gibt ein $B' \sim A$.

(2) Es gibt kein $A' \sim B$, aber es gibt ein $B' \sim A$.

(3) Es gibt ein $A' \sim B$, aber es gibt kein $B' \sim A$.

(4) Es gibt kein $A' \sim B$, und es gibt kein $B' \sim A$.

Der erste Schritt, den wir zu tun haben, besteht im Nachweis der Äquivalenz $A \sim B$ im Falle (1). Es gilt nämlich der

Satz I (Äquivalenzsatz): Zwei Mengen, von denen jede einer Teilmenge der anderen äquivalent ist, sind selbst äquivalent.

Es sei also

$$A' \sim B, \quad B' \sim A$$

und die Behauptung lautet:

$$A \sim B.$$

Vermöge der umkehrbar eindeutigen Beziehung, die zwischen B und A'

besteht, wird die Teilmenge B' von B auf eine Teilmenge A'' von A' abgebildet, so daß also

$$A \geqq A' \geqq A'', \quad A'' \sim B' \sim A.$$

Die Behauptung kann in der Form $A \sim A'$ und damit der ganze Satz in der Form ausgesprochen werden:

Wenn $A \geqq A' \geqq A''$ und $A \sim A''$, so ist auch $A \sim A'$, oder: ist A mit einer Teilmenge A'' äquivalent, so ist es auch mit jeder Teilmenge A' zwischen A und A'' äquivalent (d. h. die in A enthalten ist und A'' enthält).

Setzen wir noch

$$A'' = P, \; A' - A'' = Q, \; A - A' = R,$$

so lautet der zu beweisende Satz: ist $P + Q + R \sim P$, so ist auch $P + Q \sim P$.

Wir nehmen alle drei Mengen $P, Q, R > 0$ an, da in jedem anderen Falle der Satz trivial ist.

Bei der umkehrbar eindeutigen Abbildung zwischen $P + Q + R$

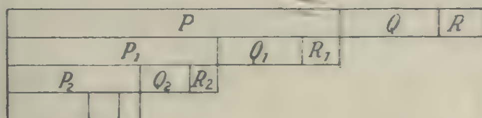


Fig. 3.

und P wird jeder Teilmenge X von $P + Q + R$ eine Teilmenge $Y = f(X)$ von P zugeordnet, und nach den Bemerkungen in Kap. II, § 4 ist

$$f(X_1 + X_2 + \ldots) = f(X_1) + f(X_2) + \cdots$$

Wir haben daher:

$$\begin{aligned} P &= f(P + Q + R) = f(P) + f(Q) + f(R) = P_1 + Q_1 + R_1 \\ P_1 &= f(P_1 + Q_1 + R_1) = f(P_1) + f(Q_1) + f(R_1) = P_2 + Q_2 + R_2 \\ P_2 &= f(P_2 + Q_2 + R_2) = f(P_2) + f(Q_2) + f(R_2) = P_3 + Q_3 + R_3 \end{aligned}$$

usw.,

wobei

$$P_1 = f(P), \; P_2 = f(P_1), \; P_3 = f(P_2), \cdots$$

(und analog für die anderen Mengen) gesetzt ist.

Hierbei ist $P > P_1 > P_2 > \ldots$. Sei nun

$$P_\omega = \mathfrak{D}(P, P_1, P_2, \ldots).$$

Ein Element p von P gehört dann entweder allen P_n $(n = 1, 2, \ldots)$ an, also zu P_ω, oder, wenn nicht, so gibt es eine kleinste Zahl n derart, daß p nicht mehr zu P_n, aber noch zu P_{n-1}, also zu

$$P_{n-1} - P_n = Q_n + R_n$$

gehört ($P_0 = P$ gesetzt). Es ist also

$$P = P_\omega + \sum_n (P_{n-1} - P_n)$$
$$= P_\omega + \sum_n (Q_n + R_n).$$

Nach dem assoziativen Gesetz und dem der Summenäquivalenz folgt daraus

$$P = P_\omega + Q_1 + R_1 + Q_2 + R_2 + Q_3 + R_3 + \cdots$$
$$\sim P_\omega + Q + R_1 + Q_1 + R_2 + Q_2 + R_3 + \cdots$$
$$= Q + P,$$

also $P \sim P + Q$, was zu beweisen war.

Dieser Beweis ist von F. Bernstein gegeben worden. Von demjenigen in der Anmerkung S. 47 bezeichneten Standpunkt aus, der auch die Theorie der endlichen Mengen in die Mengenlehre verweist, würde man einen so grundlegenden Satz vor der Einführung der natürlichen Zahlenreihe in der folgenden Art nach E. Zermelo beweisen. Nennen wir, für den Augenblick, X eine „normale" Teilmenge von $P + Q + R$, wenn sie Q und ihre Bildmenge $f(X)$ umfaßt, welche Bedingung wegen $f(X) \leqq P$ durch

$$X \geqq f(X) + Q$$

ausgedrückt wird. Man sieht unmittelbar, daß $X' = f(X) + Q$ wieder eine normale Teilmenge ist, daß der Durchschnitt beliebig vieler normaler Teilmengen wieder normal ist, und daß es normale Teilmengen gibt, z. B. $P + Q + R$. Ist jetzt X der Durchschnitt aller normalen Teilmengen, oder die kleinste normale Teilmenge, so muß nach dem Gesagten

$$X = f(X) + Q$$

sein, und setzt man mit Rücksicht auf $f(X) \leqq P$

$$P = Y + f(X),$$

so ist

$$P + Q = Y + f(X) + Q = Y + X$$
$$\sim Y + f(X) = P.$$

Um beide Beweise in Zusammenhang zu bringen, genügt die Bemerkung, daß eine normale Teilmenge ihre eigene Bildmenge, also auch die jeder ihrer Teilmengen enthalten soll; da sie Q enthalten soll, enthält sie auch $f(Q) = Q_1$, $f(Q_1) = Q_2$ usw. und die kleinste normale Teilmenge ist

$$X = Q + Q_1 + Q_2 + \ldots$$

Nach dem Beweise des Äquivalenzsatzes kehren wir zu der obigen vierfachen Disjunktion zurück.

Im Falle (1) ist also $\mathfrak{a} = \mathfrak{b}$.

Im Falle (2) definieren wir $\mathfrak{a} < \mathfrak{b}$.

Im Falle (3) definieren wir $\mathfrak{a} > \mathfrak{b}$.

Hiermit hätten wir zwar erreicht, daß diese drei Beziehungen einander ausschließen, also höchstens eine von ihnen eintreten kann; aber es ist noch nicht erreicht, daß von ihnen wirklich eine eintreten muß, denn es ist mit der Möglichkeit des vierten Falles zu rechnen, in welchem uns nichts übrig bliebe, als die beiden Mengen A, B und ihre Mächtigkeiten für unvergleichbar zu erklären (was wir vorläufig durch das Zeichen $\mathfrak{a} \parallel \mathfrak{b}$ ausdrücken wollen). Bei den endlichen Mengen ist der vierte Fall ausgeschlossen, da die Bildung von Paaren eben zu einem Ende kommen muß und die eine Menge dann auf eine Teilmenge der anderen eineindeutig bezogen ist. Erst später, in der Lehre von den wohlgeordneten Mengen, werden wir allgemein zeigen können, daß der vierte Fall überhaupt ausgeschlossen ist und die vierfache Disjunktion also in eine dreifache, in die Trichotomie

$$\mathfrak{a} = \mathfrak{b} \text{ oder } \mathfrak{a} < \mathfrak{b} \text{ oder } \mathfrak{a} > \mathfrak{b}$$

übergeht. Vorläufig können wir uns das höchstens in einer undeutlichen Weise plausibel machen, indem wir sagen: man nehme ein Element a_1 und ein Element b_1, die zu einem Paar p_1 vereinigt werden, dann ein zweites Element a_2 und ein zweites Element b_2, ein drittes Element a_3 und ein drittes Element b_3 usw. Kommt man hierbei nicht zu Ende, kann man also für jede natürliche Zahl n ein Paar $p_n = (a_n, b_n)$ bilden, ohne eine der Mengen zu erschöpfen, so muß man das Verfahren weiter fortsetzen. Aber eben diese Fortsetzung über die Folge der natürlichen Zahlen hinaus, bis zur Aufzehrung einer der beiden Mengen, ist hier noch eine ganz nebelhafte Forderung und muß in einer erst später zu präzisierenden Anordnung geschehen.

Ist A einer Teilmenge von B äquivalent, tritt also der Fall (1) oder (2) ein, so schreiben wir

$$\mathfrak{a} \leqq \mathfrak{b},$$

welche Formel also bedeutet, daß entweder $\mathfrak{a} = \mathfrak{b}$ oder $\mathfrak{a} < \mathfrak{b}$.

Ist $\quad \mathfrak{a} = \mathfrak{b},\ \mathfrak{a} < \mathfrak{b},\ \mathfrak{a} > \mathfrak{b},\ (\mathfrak{a} \parallel \mathfrak{b}),$

so ist $\quad \mathfrak{b} = \mathfrak{a},\ \mathfrak{b} > \mathfrak{a},\ \mathfrak{b} < \mathfrak{a},\ (\mathfrak{b} \parallel \mathfrak{a}).$

Aus $\mathfrak{a} = \mathfrak{b}$, $\mathfrak{b}\,\varrho\,\mathfrak{c}$ folgt $\mathfrak{a}\,\varrho\,\mathfrak{c}$, wenn ϱ eine der vier Relationen ist.

Aus $\mathfrak{a} < \mathfrak{b}$, $\mathfrak{b} < \mathfrak{c}$ folgt $\mathfrak{a} < \mathfrak{c}$ (transitives Gesetz).

§ 3. Summe, Produkt, Potenz.

Wir hatten gesehen, daß $\mathfrak{S}(A, B)$ und $\mathfrak{D}(A, B)$ bei Ersetzung von A, B durch äquivalente Mengen ihrerseits nicht in äquivalente Mengen überzugehen brauchen. Wohl aber ist $A + B \sim A' + B'$, falls $A \sim A'$, $B \sim B'$ und A mit B, A' mit B' keine gemeinsamen Elemente hat.

Wir definieren daher die Summe $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$ zweier Kardinalzahlen $\mathfrak{a}$, $\mathfrak{b}$ als Mächtigkeit einer Menge $A + B$, wenn A und B irgend zwei fremde Mengen von den Mächtigkeiten $\mathfrak{a}$ und $\mathfrak{b}$ sind. Nach dem eben bemerkten ist $\mathfrak{a} + \mathfrak{b}$ nur von $\mathfrak{a}$ und $\mathfrak{b}$, nicht von den gewählten Mengen A und B abhängig.

Allgemein: sind den Elementen i einer Menge J Kardinalzahlen $\mathfrak{a}_i$ zugeordnet (die also einen Komplex mit dem Argument J bilden), so verstehen wir unter der Summe

$$\sum_i^J \mathfrak{a}_i = \mathfrak{a}_i + \mathfrak{a}_k + \mathfrak{a}_l + \dots$$

die Mächtigkeit der Mengensumme

$$\sum_i^J A_i = A_i + A_k + A_l + \dots,$$

wo die A_i paarweise fremde Mengen von den Mächtigkeiten $\mathfrak{a}_i$ sind. Die Zahlensumme hängt von der Wahl dieser Mengen nicht ab, da die Mengensumme bei Ersetzung der A_i durch äquivalente, wieder paarweise fremde Mengen in eine äquivalente Summe übergeht. Das kommutative und assoziative Gesetz ist aus Kap. II, § 3 unmittelbar zu entnehmen.

Beispiele. In Kap. I, § 8 sahen wir, daß eine Folge mit einer endlichen Menge oder einer Folge sich wieder zu einer Folge vereinigen läßt. Das gibt also die Gleichungen (n eine natürliche Zahl)

$$\aleph_0 + n = \aleph_0 + \aleph_0 = \aleph_0,$$

die ersten Beispiele jener in Aussicht gestellten Abweichungen zwischen endlichen und unendlichen Zahlen. Das unendliche $\aleph_0$ wird also durch Hinzufügung des endlichen n gar nicht vermehrt; aber selbst durch Hinzufügung seiner selbst wird es nicht vergrößert. Aus den weiteren Überlegungen des zitierten Paragraphen oder unter Anwendung des assoziativen Gesetzes finden wir weiter

$$\aleph_0 + \aleph_0 + \aleph_0 = (\aleph_0 + \aleph_0) + \aleph_0 = \aleph_0 + \aleph_0 = \aleph_0$$

und allgemein

$$\sum_i^{(1, \dots, n)} \aleph_0 = \aleph_0 + \aleph_0 + \dots + \aleph_0 = \aleph_0.$$

Aber selbst eine Folge von Folgen oder endlichen Mengen > 0 gibt durch Vereinigung wieder eine Folge. Ordnen wir also jeder natürlichen Zahl i eine Kardinalzahl $\mathfrak{a}_i$ zu, die entweder eine natürliche Zahl oder $= \aleph_0$ ist, so ist

$$\mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \dots = \aleph_0,$$

z. B.

$$1 + 1 + 1 + \dots = \aleph_0,$$
$$2 + 2 + 2 + \dots = \aleph_0,$$
$$1 + 2 + 3 + \dots = \aleph_0,$$
$$\aleph_0 + \aleph_0 + \aleph_0 + \dots = \aleph_0.$$

Als **Produkt**

$$\mathfrak{P}_i^J \mathfrak{a}_i = \mathfrak{a}_i \mathfrak{a}_k \mathfrak{a}_l \ldots$$

definieren wir die Mächtigkeit des Mengenprodukts

$$\mathfrak{P}_i^J A_i = A_i A_k A_l \ldots,$$

worin A_i eine Menge der Mächtigkeit $\mathfrak{a}_i$ ist. Wiederum ist die Wahl dieser Mengen gleichgültig.

Insbesondere ist $\mathfrak{a}\mathfrak{b} = \mathfrak{b}\mathfrak{a}$ die Mächtigkeit jedes Produkts[1] AB oder der Menge der geordneten Paare (a, b).

Die Summe $\sum_i^J \mathfrak{a}_i$ geht, wenn alle $\mathfrak{a}_i = \mathfrak{a}$, in das Produkt $\mathfrak{a}\mathfrak{i}$ über ($\mathfrak{i}$ die Mächtigkeit von J).

Wir begnügen uns, bezüglich dieser Tatsachen sowie bezüglich des kommutativen, assoziativen und distributiven Gesetzes (zwischen Summe und Produkt) an die entsprechenden Formeln in Kap. II, § 3 zu erinnern.

Als **Potenz** $\mathfrak{a}^{\mathfrak{i}}$ definieren wir die Mächtigkeit von A^J, wenn A, J von der Mächtigkeit $\mathfrak{a}$, $\mathfrak{i}$ sind — wobei die Wahl dieser Mengen wieder gleichgültig ist.

Das Produkt $\mathfrak{P}_i^J \mathfrak{a}_i$ geht, wenn alle $\mathfrak{a}_i = \mathfrak{a}$, in die Potenz $\mathfrak{a}^{\mathfrak{i}}$ über.

Aus dem mehrfach zitierten Kap. II, § 3 entnehmen wir die Potenzgesetze:

$$\mathfrak{a}^{\mathfrak{b}_1} \cdot \mathfrak{a}^{\mathfrak{b}_2} \ldots = \mathfrak{a}^{\mathfrak{b}_1 + \mathfrak{b}_2 + \cdots},$$

$$(\mathfrak{a}^{\mathfrak{b}})^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}\mathfrak{c}},$$

$$\mathfrak{a}_1^{\mathfrak{b}} \cdot \mathfrak{a}_2^{\mathfrak{b}} \ldots = (\mathfrak{a}_1 \mathfrak{a}_2 \ldots)^{\mathfrak{b}}.$$

Beispiele.

$$2\aleph_0 = \aleph_0 + \aleph_0 = \aleph_0,$$

$$3\aleph_0 = \aleph_0 + \aleph_0 + \aleph_0 = \aleph_0,$$

$$n\aleph_0 = \aleph_0 + \aleph_0 + \ldots + \aleph_0 = \aleph_0.$$

Die Verwandlung einer Doppelfolge in eine einfache gibt

$$\aleph_0 \aleph_0 = \aleph_0 + \aleph_0 + \aleph_0 + \ldots = \aleph_0,$$

also

$$\aleph_0^2 = \aleph_0 \aleph_0 = \aleph_0,$$

ebenso

$$\aleph_0^3 = \aleph_0 \aleph_0 \aleph_0 = \aleph_0,$$

$$\aleph_0^n = \aleph_0 \aleph_0 \ldots \aleph_0 = \aleph_0.$$

Durchläuft i die natürlichen Zahlen, so ist

$$\mathfrak{P}_i \mathfrak{a}_i = \mathfrak{a}_1 \mathfrak{a}_2 \mathfrak{a}_3 \ldots$$

---

[1] mit beliebigem, aus zwei Elementen bestehenden Argument.

die Mächtigkeit der Menge aller Elementenkomplexe oder Folgen $(a_1, a_2, a_3, \ldots)$, wo jedes a_i eine Menge A_i der Mächtigkeit $\mathfrak{a}_i$ durchläuft; sind alle diese Mächtigkeiten gleich $(\mathfrak{a}_i = \mathfrak{a})$, so entsteht die Potenz

$$\mathfrak{P}_i \mathfrak{a} = \mathfrak{a}\mathfrak{a}\mathfrak{a} \ldots = \mathfrak{a}^{\aleph_0} .$$

Z. B. ist $2^{\aleph_0}$ die Mächtigkeit der Menge aller Folgen, die aus zwei Elementen, etwa den Ziffern 0, 1 gebildet sind; $10^{\aleph_0}$ die der Menge aller Folgen, die aus 10 Elementen, etwa den Ziffern 0, 1, ..., 9 gebildet sind, oder, wenn man jeder solchen Folge $(a_1, a_2, a_3, \ldots)$ den Dezimalbruch

$$0, a_1\, a_2\, a_3 \ldots = \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3} + \cdots$$

zuordnet, so ist $10^{\aleph_0}$ die Mächtigkeit der Menge dieser Dezimalbrüche, wobei zwei Dezimalbrüche auch dann noch als verschieden anzusehen sind, wenn sie dieselbe reelle Zahl darstellen (z. B. 0,4999 ... und 0,5000 ...).

Die Mächtigkeit der Menge aller Folgen natürlicher Zahlen ist

$$\aleph_0^{\aleph_0} = \aleph_0\, \aleph_0\, \aleph_0 \ldots$$

Hat A die Mächtigkeit $\mathfrak{a}$, so hat die Menge aller Teilmengen von A die Mächtigkeit $2^{\mathfrak{a}}$ (Kap. II, § 2, Schluß).

§ 4. Ungleichungen zwischen Mächtigkeiten.

Unsere Kenntnisse in bezug auf die Vergleichung von Kardinalzahlen sind noch sehr gering; wir stellen hier die bezüglichen Sätze zusammen.

Zunächst folgt aus der Definition von Summe und Produkt unmittelbar, daß bei gleichem Argument

$$\sum_i \mathfrak{a}_i \leqq \sum_i \mathfrak{b}_i ,$$

$$\mathfrak{P}_i \mathfrak{a}_i \leqq \mathfrak{P}_i \mathfrak{b}_i$$

ist, falls, für jedes i, $\mathfrak{a}_i \leqq \mathfrak{b}_i$ ist. Selbst wenn, für jedes i, $\mathfrak{a}_i < \mathfrak{b}_i$ ist, darf in diesen Formeln das Gleichheitszeichen nicht ausgeschlossen werden,[1] wie für die Summen das Beispiel $(i = 1, 2, 3, \ldots)$

$$\sum_i 1 = \sum_i 2 = \aleph_0$$

lehrt; wir werden nachher sehen, daß auch $\mathfrak{P}_i 2 = \mathfrak{P}_i 3$ ist (§ 5, (1)).

Insbesondere ist

$$\mathfrak{a} = \mathfrak{a} + 0 \leqq \mathfrak{a} + \mathfrak{b} ,$$

$$\mathfrak{a} = \mathfrak{a} \,.\, 1 \leqq \mathfrak{a}\mathfrak{b} \quad \text{für} \quad \mathfrak{b} > 0 .$$

Ein Produkt verschwindet nur, wenn einer seiner Faktoren verschwindet. Denn der Mengenkomplex $(A_1, A_2, \ldots)$ enthält, wenn alle $A_i > 0$, mindestens einen Elementenkomplex $(a_1, a_2, \ldots)$.

---

[1] Vgl. dagegen den Satz IV.

Mit Hilfe des assoziativen Gesetzes kann man auch Summen und Produkte verschiedener Argumente vergleichen. Z. B. ist

$$\sum_i^{J_1+J_2} \mathfrak{a}_i = \sum_i^{J_1} \mathfrak{a}_i + \sum_i^{J_2} \mathfrak{a}_i \geqq \sum_i^{J_1} \mathfrak{a}_i,$$

$$\mathfrak{P}_i^{J_1+J_2} \mathfrak{a}_i = \mathfrak{P}_i^{J_1} \mathfrak{a}_i \cdot \mathfrak{P}_i^{J_2} \mathfrak{a}_i \geqq \mathfrak{P}_i^{J_1} \mathfrak{a}_i,$$

letzteres nur, falls, für jedes Element i von J_2, $\mathfrak{a}_i > 0$.

Für Potenzen folgt daraus:

$$\mathfrak{a}^{\mathfrak{i}} \leqq (\mathfrak{a}+\mathfrak{b})^{\mathfrak{i}}, \quad \mathfrak{a}^{\mathfrak{i}} \leqq \mathfrak{a}^{\mathfrak{i}+\mathfrak{k}}.$$

Weiter beweisen wir jetzt folgende Sätze:

I. Jede unendliche Mächtigkeit ist $\geqq \aleph_0$; d. h. jede unendliche Menge enthält eine abzählbare Teilmenge.

Denn die unendliche Menge A enthält mindestens ein Element a_1; die Menge $A_1 = A - \{a_1\}$ ist noch unendlich und enthält wieder ein Element a_2, das von a_1 verschieden ist; die Menge

$$A_2 = A_1 - \{a_2\} = A - \{a_1, a_2\}$$

enthält wieder ein von a_1, a_2 verschiedenes Element a_3 usw. Da für jede natürliche Zahl n die Menge $A_n = A - \{a_1, a_2, \ldots, a_n\}$ noch unendlich ist und ein weiteres Element a_{n+1} enthält, so enthält A die Menge $\{a_1, a_2, \ldots\}$, die mit der Menge $\{1, 2, \ldots\}$ der natürlichen Zahlen äquivalent ist.[1]

Mit Rücksicht auf I kann $\aleph_0$ als kleinste unendliche Mächtigkeit bezeichnet werden; man nennt es auch die erste Mächtigkeit.

Jede unendliche Teilmenge einer abzählbaren Menge (z. B. die Menge der Primzahlen) ist abzählbar. Denn ihre Mächtigkeit ist $\leqq \aleph_0$, nach I aber auch $\geqq \aleph_0$, also nach dem Äquivalenzsatz $= \aleph_0$.

Jede unendliche Menge A kann nach dem Obigen in der Gestalt

$$A = B + C$$

dargestellt werden, wo $C = \{c_1, c_2, \ldots\}$ eine abzählbare Menge ist. Oder jede unendliche Mächtigkeit $\mathfrak{a}$ kann in der Form

$$\mathfrak{a} = \mathfrak{b} + \aleph_0$$

dargestellt werden. Es folgt daraus für jede natürliche Zahl n

$$\mathfrak{a} + n = (\mathfrak{b} + \aleph_0) + n = \mathfrak{b} + (\aleph_0 + n) = \mathfrak{b} + \aleph_0 = \mathfrak{a},$$

ebenso

$$\mathfrak{a} + \aleph_0 = (\mathfrak{b} + \aleph_0) + \aleph_0 = \mathfrak{b} + (\aleph_0 + \aleph_0) = \mathfrak{b} + \aleph_0 = \mathfrak{a};$$

jede unendliche Kardinalzahl bleibt also bei Hinzufügung einer endlichen oder von $\aleph_0$ ungeändert.

---

[1] Hierzu ist der spätere Beweis (Kap. V) zu vergleichen, daß jede Menge wohlgeordnet werden kann.

Jede unendliche Menge ist einer echten Teilmenge äquivalent. Z. B. ist ja bei der obigen Zerlegung

$$A = B + \{c_1, c_2, c_3, \ldots\}$$
$$\sim B + \{c_2, c_3, c_4, \ldots\}$$
$$\sim B + \{c_2, c_4, c_6, \ldots\}$$

und dergl. Umgekehrt ist eine Menge, die einer echten Teilmenge äquivalent ist, notwendig unendlich, d. h. nicht endlich. Nach dem Vorgang von R. Dedekind wird die Äquivalenz mit einer echten Teilmenge bisweilen als Definition der unendlichen Mengen gewählt.

Entfernt man von einer unendlichen Menge A endlich viele Elemente, so bleibt eine mit A äquivalente Menge. Denn ist

$$\mathfrak{a} = \mathfrak{b} + n,$$

so ist mit $\mathfrak{a}$ auch $\mathfrak{b}$ unendlich, also nach dem vorigen

$$\mathfrak{b} = \mathfrak{b} + n = \mathfrak{a}.$$

Entfernt man von einer unabzählbaren[1] Menge A eine abzählbare Menge, so bleibt eine mit A äquivalente Menge. Denn ist

$$\mathfrak{a} = \mathfrak{b} + \aleph_0,$$

so ist unter den gemachten Voraussetzungen $\mathfrak{b}$ unendlich, also

$$\mathfrak{b} = \mathfrak{b} + \aleph_0 = \mathfrak{a}.$$

II. Es ist stets $2^{\mathfrak{a}} > \mathfrak{a}$, d. h. die Menge der Teilmengen einer Menge A ist von größerer Mächtigkeit als A selbst.

Bezeichnet $\mathfrak{A}$ das System der Teilmengen von A, so ist zunächst A einem Teilsystem von $\mathfrak{A}$ äquivalent, denn jedem Element a entspricht umkehrbar eindeutig die aus diesem einen Element a bestehende Teilmenge $\{a\}$. Jedenfalls ist also $\mathfrak{a} \leqq 2^{\mathfrak{a}}$.

Daß aber nicht $\mathfrak{a} = 2^{\mathfrak{a}}$ sein kann (also $\mathfrak{a} < 2^{\mathfrak{a}}$ sein muß), zeigen wir so: ist $\mathfrak{A}'$ ein mit A äquivalentes System von Teilmengen ($\mathfrak{A}' \leqq \mathfrak{A}$, $\mathfrak{A}' \sim A$), so kann $\mathfrak{A}'$ nicht mit dem ganzen System $\mathfrak{A}$ identisch sein. In der Tat, es entspreche umkehrbar eindeutig jedem Element a eine Teilmenge A_a, und die letzteren bilden also das System $\mathfrak{A}'$. Die Menge

$$B = A - \sum_a^A \mathfrak{D}(\{a\}, A_a)$$

ist dann eine Teilmenge von A und besteht aus denjenigen Elementen a, die nicht Elemente der zugehörigen Menge A_a sind; d. h. gehört a zu A_a, so gehört es nicht zu B, und gehört a nicht zu A_a, so gehört es zu B. Infolgedessen ist B von jedem A_a verschieden und gehört dem System $\mathfrak{A}'$ nicht an.

Hiermit erkennen wir, daß nicht etwa alle unendlichen Mengen

[1] Darunter verstehen wir eine unendliche, nicht abzählbare Menge.

von gleicher Mächtigkeit sind; das Unendliche ist nicht schlechthin die indifferente Negation des Endlichen, sondern wie dieses der Abstufung fähig. Wir erhalten aus einer unendlichen Mächtigkeit $\mathfrak{a}$ (z. B. $\aleph_0$) der Reihe nach höhere, indem wir bilden

$$\mathfrak{a}_1 = 2^{\mathfrak{a}} > \mathfrak{a},$$
$$\mathfrak{a}_2 = 2^{\mathfrak{a}_1} > \mathfrak{a}_1,$$
$$\mathfrak{a}_3 = 2^{\mathfrak{a}_2} > \mathfrak{a}_2 \text{ usw.}$$

Übrigens gilt der Satz auch für endliche Kardinalzahlen, z. B.

$$1 = 2^0 > 0,$$
$$2 = 2^1 > 1,$$
$$4 = 2^2 > 2,$$
$$8 = 2^3 > 3 \text{ usw.}$$

III. Ist $(\mathfrak{a}_i, \mathfrak{a}_k, \ldots)$ ein Komplex von Kardinalzahlen, worin zu jeder Kardinalzahl $\mathfrak{a}_i$ eine größere $\mathfrak{a}_k > \mathfrak{a}_i$ vorhanden ist, so ist die Summe

$$\sum_i \mathfrak{a}_i$$

größer als jede Zahl des Komplexes.

In der Tat ist ja

$$\sum_i \mathfrak{a}_i \geqq \mathfrak{a}_k > \mathfrak{a}_i.$$

Das gleiche gilt von einer Menge von Kardinalzahlen; eine solche Menge umfaßt also niemals alle Kardinalzahlen, und die Gesamtheit aller Kardinalzahlen kann nicht als Menge angesehen werden (vgl. Kap. I, § 1).

Danach kommt der Prozeß des Aufsteigens zu immer höheren Mächtigkeiten nie zum Abschluß. Haben wir z. B. aus einer Mächtigkeit $\mathfrak{a}$ nach der vorhin angegebenen Methode die Folge aufsteigender Mächtigkeiten

$$\mathfrak{a} < \mathfrak{a}_1 < \mathfrak{a}_2 < \cdots$$

gebildet, so ist

$$\mathfrak{b} = \mathfrak{a} + \mathfrak{a}_1 + \mathfrak{a}_2 + \cdots$$

eine Mächtigkeit über allen $\mathfrak{a}_i$, mit der nun wieder $2^{\mathfrak{b}} > \mathfrak{b}$ usf. gebildet werden kann.

IV. (Satz von J. König.) Sind $(\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_i, \ldots)$ und $(\mathfrak{b}_1, \mathfrak{b}_2, \ldots, \mathfrak{b}_i, \ldots)$ zwei Zahlenkomplexe desselben Arguments und ist, für jedes i, $\mathfrak{a}_i < \mathfrak{b}_i$, so ist

$$\sum_i \mathfrak{a}_i < \mathfrak{P}_i \mathfrak{b}_i.$$

Beweis. Die Summe ist die Mächtigkeit einer Menge

$$S = \sum_i A_i = A_1 + A_2 + \cdots + A_i + \cdots,$$

wo die A_i paarweise fremde Mengen der Mächtigkeit $\mathfrak{a}_i$ sind, das Produkt die Mächtigkeit einer Menge

$$P = \mathfrak{P}_i B_i = B_1 B_2 \ldots B_i \ldots,$$

wo die B_i Mengen der Mächtigkeit $\mathfrak{b}_i$ sind. A_i ist hierbei einer Teilmenge von B_i äquivalent, und wegen der Ersetzbarkeit der A_i oder B_i durch äquivalente Mengen können wir $A_i < B_i$,

$$B_i = A_i + C_i$$

annehmen, wo $C_i > 0$, weil sonst $\mathfrak{a}_i = \mathfrak{b}_i$ sein würde.

Die Elemente von P sind die verschiedenen Elementkomplexe

$$p = (b_1, b_2, \ldots, b_i, \ldots),$$

wo $\mathfrak{b}_i$ die Menge B_i zu durchlaufen hat.

Zunächst ist nun S einer Teilmenge von P äquivalent, also

$$\sum_i \mathfrak{a}_i \leqq \mathfrak{P}_i \mathfrak{b}_i.$$

Dann versteht man unter c_i ein bestimmtes, festes Element von C_i, während a_i die sämtlichen Elemente von A_i durchlaufen soll, so bilden die Komplexe

$$\begin{array}{c}(a_1, c_2, c_3, \ldots, c_i, \ldots)\\ (c_1, a_2, c_3, \ldots, c_i, \ldots)\\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot\\ (c_1, c_2, c_3, \ldots, a_i, \ldots)\\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot\end{array}$$

Mengen, die resp. mit $A_1, A_2, \ldots, A_i, \ldots$ äquivalent und paarweise fremd sind, da ja z. B. wegen $a_1 \neq c_1$, $a_2 \neq c_2$ kein Komplex der ersten Zeile mit einem der zweiten Zeile übereinstimmen kann. Die Menge aller dieser Komplexe ist also eine mit

$$A_1 + A_2 + \ldots + A_i + \ldots = S$$

äquivalente Teilmenge von P.

Daß andererseits nicht $S \sim P$ sein kann, zeigen wir wieder so: Q sei eine mit S äquivalente Teilmenge von P $(Q \leqq P, Q \sim S)$, dann kann Q nicht mit der ganzen Menge P identisch sein. Vermöge der umkehrbar eindeutigen Beziehung zwischen S und Q entspricht jedem A_i eine äquivalente Teilmenge Q_i von Q, und es ist

$$Q = Q_1 + Q_2 + \cdots + Q_i + \cdots$$

Es sei $p = (b_1, b_2, \ldots, b_i, \ldots)$ ein zu Q_i gehöriger Komplex, und wir fassen insbesondere das dem Index i entsprechende Element b_i ins Auge. Wegen $A_i \sim Q_i$ entspricht jedem Element a_i von A_i ein solches Element b_i, jedem Element a_i' von A_i wieder ein Element b_i', das von b_i nicht notwendig verschieden ist. Diese sämtlichen Elemente b_i bilden also, wenn p die Menge Q_i durchläuft, eine Menge D_i

$\leqq B_i$), deren Elemente eindeutig, nicht notwendig umkehrbar eindeutig, den Elementen a_i entsprechen, und deren Mächtigkeit also $\mathfrak{b}_i \leqq \mathfrak{a}_i$. Setzt man also jetzt

$$B_i = D_i + E_i,$$

so ist $E_i > 0$, da sonst $\mathfrak{b}_i \leqq \mathfrak{a}_i$ wäre. Wählt man nun endlich aus jeder Menge E_i beliebig ein Element e_i, so ist

$$(e_1, e_2, \ldots, e_i, \ldots)$$

ein Elementenkomplex, der nicht zu Q_i gehört, da ein zu Q_i gehöriger Komplex von der Form

$$(b_1, b_2, \ldots, d_i, \ldots)$$

und $d_i \neq e_i$ ist. Da dies für jedes i gilt, so gehört jener Komplex auch nicht zu $Q = \sum_i Q_i$, und Q ist demnach nicht mit P identisch.

Man kann den Satz II als Spezialfall dieses Satzes IV auffassen. In der Tat ist ja

$$\mathfrak{a} = \sum_a^A 1 < \mathfrak{P}_a^A 2 = 2^{\mathfrak{a}},$$

z. B.

$$\aleph_0 = 1 + 1 + 1 + \ldots < 2 . 2 . 2 . \ldots = 2^{\aleph_0}.$$

Eine andere Anwendung bezieht sich auf die Summe $\mathfrak{a} = \mathfrak{a}_1 + \mathfrak{a}_2 + \ldots$ und das Produkt $\mathfrak{b} = \mathfrak{a}_1 \mathfrak{a}_2 \ldots$ einer aufsteigenden Folge von Mächtigkeiten, wo

$$0 < \mathfrak{a}_1 < \mathfrak{a}_2 < \ldots .$$

Man hat hier

$$\mathfrak{a}_n < \mathfrak{a}_{n+1} < \mathfrak{a},$$

also

$$\mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \ldots < \mathfrak{a}_2 \mathfrak{a}_3 \mathfrak{a}_4 \ldots = 1 . \mathfrak{a}_2 \mathfrak{a}_3 \mathfrak{a}_4 \ldots$$
$$\leqq \mathfrak{a}_1 \mathfrak{a}_2 \mathfrak{a}_3 \ldots \leqq \mathfrak{a}\mathfrak{a}\mathfrak{a} \ldots,$$

d. h.

$$\mathfrak{a} < \mathfrak{b} \leqq \mathfrak{a}^{\aleph_0}.$$

Die Summe ist also kleiner als das Produkt, dieses höchstens gleich der $\aleph_0$ten Potenz der Summe.

§ 5. Die Mächtigkeiten $\aleph_0$, $2^{\aleph_0}$, $2^{2^{\aleph_0}}$.

Von diesen wichtigsten Mächtigkeiten sind noch einige Worte zu sagen, wobei sich, da wir sie bereits als Beispiele zu allgemeinen Tatsachen zu verwenden hatten, einige Wiederholungen nicht vermeiden lassen.

Die erste Mächtigkeit $\aleph_0$, die der abzählbaren Mengen, kommt allen Mengen zu, die sich als Folgen darstellen lassen, so auch der Summe von zwei, drei oder n Folgen, einer Folge von Folgen oder Doppelfolge oder dem Produkt von zwei Folgen, dem Produkt

von drei oder n Folgen. Wir zitieren noch einige Beispiele außer den genannten.

Die Menge der ganzen Zahlen

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

ist abzählbar, da sie sich in Gestalt der Folge

$$0, 1, -1, 2, -2, 3, -3, \ldots$$

bringen läßt. Oder auch: sie setzt sich aus der Menge der negativen ganzen Zahlen (die der Menge der positiven äquivalent ist), der Zahl 0 und der der positiven ganzen Zahlen zusammen, hat also die Mächtigkeit

$$\aleph_0 + 1 + \aleph_0 = \aleph_0.$$

Die Menge der rationalen Zahlen ist abzählbar. Betrachten wir etwa die positiven rationalen Zahlen p/q (p, q natürliche Zahlen ohne gemeinsamen Teiler); sie lassen sich in Gestalt einer Folge schreiben, indem man etwa zunächst nach der Summe von Zähler und Nenner, sodann nach dem Zähler ordnet, d. h. p/q vor p'/q' setzt, wenn entweder $p+q<p'+q'$ oder $p+q=p'+q'$, $p<p'$. So entsteht die Folge

$$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{1}, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1}, \frac{1}{5}, \frac{5}{1}, \ldots$$

Wir können auch sagen: lassen wir jeder rationalen Zahl p/q das geordnete Zahlenpaar (p, q) entsprechen, so erhalten wir eine unendliche Menge von Zahlenpaaren (nicht alle, weil ja p und q teilerfremd sein sollen). Da die Menge aller Paare die Mächtigkeit $\aleph_0 \aleph_0 = \aleph_0$ hat, so hat jede unendliche Teilmenge von ihr ebenfalls die Mächtigkeit $\aleph_0$.

Die Menge aller rationalen Zahlen besteht aus den positiven, den negativen und der Zahl 0, hat also wieder die Mächtigkeit $\aleph_0$.

Auch die Menge aller rationalen Zahlen r eines Intervalls ($a \leqq r \leqq b$, für $a<b$) ist abzählbar, da sie unendlich ist. Wir bringen z. B. die rationalen Zahlen zwischen 0 und 1 noch auf eine zweite Art in eine Folge, indem wir erst nach den Nennern, sodann nach den Zählern ordnen:

$$\frac{0}{1}, \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \ldots$$

Die Äquivalenz der Menge der ganzen Zahlen mit der doch viel umfassenderen der rationalen Zahlen gehört mit zu den Tatsachen der Mengenlehre, die bei erster Bekanntschaft den Eindruck des Erstaunlichen, ja Paradoxen hervorrufen: namentlich wenn man das geometrische Bild (die Zuordnung zwischen Zahlen und Punkten einer geraden Linie) vor Augen hat und sich einerseits die in endlichen Abständen isoliert liegenden „ganzzahligen" Punkte, anderer-

seits die über die ganze Linie wie ein Staub von mehr als mikroskopischer Feinheit verteilten „rationalen“ Punkte vergegenwärtigt.

Die Menge der algebraischen Zahlen ist abzählbar, die ihrerseits wieder unendlich umfassender ist als die der rationalen Zahlen. Die (reelle oder komplexe) algebraische Zahl x genüge der Gleichung $n$ten Grades

$$(\alpha) \qquad x^n + r_1 x^{n-1} + r_2 x^{n-2} + \ldots + r_{n-1} x + r_n = 0,$$

wo n eine natürliche Zahl und $r_1, r_2, \ldots, r_n$ rationale Zahlen sind. Sie heißt eine algebraische Zahl $n$ten Grades, wenn sie einer Gleichung $n$ten und nicht niedrigeren Grades genügt; es kann nur eine solche Gleichung niedrigsten Grades von der obigen Form geben, und diese Gleichung ist irreduzibel, d. h. ihre linke Seite nicht in ein Produkt gleichgeformter Ausdrücke niedrigeren Grades zerlegbar. Beschränken wir uns also auf irreduzible Gleichungen, so gehört zu jeder algebraischen Zahl x eine einzige Gleichung (α) und zu jeder Gleichung (α) gehören n algebraische Zahlen, ihre n verschiedenen Wurzeln. Ordnen wir nun der Gleichung (α) den Komplex rationaler Zahlen $(r_1, r_2, \ldots, r_n)$ zu. Es gibt im ganzen $\aleph_0{}^n = \aleph_0$ solcher Komplexe, darunter unendlich viele, also wieder $\aleph_0$, die irreduziblen Gleichungen entsprechen; d. h. es gibt $\aleph_0$ irreduzible Gleichungen $n$ten Grades und $n \cdot \aleph_0 = \aleph_0$ algebraische Zahlen $n$ten Grades, also endlich

$$\sum_n \aleph_0 = \aleph_0 + \aleph_0 + \aleph_0 + \ldots = \aleph_0$$

algebraische Zahlen überhaupt. Die rationalen Zahlen sind die algebraischen Zahlen ersten Grades.

Die Menge aller endlichen Komplexe, die aus den Elementen einer endlichen oder abzählbaren Menge A gebildet werden können, ist abzählbar. Denn die Mächtigkeit der Menge der n-gliedrigen Komplexe $(a_1, a_2, \ldots, a_n)$ ist $\mathfrak{a}^n$, wenn $\mathfrak{a}$ die Mächtigkeit von A, also eine natürliche Zahl oder $\aleph_0$ ist, und die der Menge aller Komplexe

$$\sum_n \mathfrak{a}^n = \aleph_0 .$$

Man sieht, wie man dies auf die vorigen Beispiele anwenden kann; wenn man z. B. die n Wurzeln der irreduziblen Gleichung (α) irgendwie numeriert, so kann man ihrer $k$ten Wurzel den $(n+1)$ gliedrigen Komplex $(r_1, r_2, \ldots, r_n, k)$ zuordnen. Auch auf außermathematische Dinge ist diese Betrachtung häufig übertragen worden. Aus einem „Alphabet“, d. h. einer endlichen Menge von „Buchstaben“, kann man eine abzählbare Menge endlicher Buchstabenkomplexe, d. h. „Worte“ bilden, unter denen sich natürlich auch sinnlose wie abracadabra befinden. Nimmt man zu den Buchstaben weitere Elemente wie

Interpunktionszeichen, Druckspatien, Ziffern, Noten usw. hinzu, so sieht man, daß auch die Menge aller Bücher, Kataloge, Symphonien, Opern abzählbar ist und abzählbar bleiben würde, wenn man selbst abzählbar viele Zeichen (aber für jeden Komplex nur endlich viele) verwenden wollte. Beschränkt man dagegen, bei endlicher Zeichenzahl, die Komplexe auf eine Maximalzahl von Elementen, indem man etwa Worte von mehr als hundert Buchstaben und Bücher von mehr als einer Million Worten für unstatthaft erklärt, so werden diese Mengen endlich, und wenn man mit Giordano Bruno eine unendliche Menge von Weltkörpern annimmt, mit sprechenden, schreibenden und musizierenden Bewohnern, so folgt mit mathematischer Gewißheit, daß auf unendlich vielen dieser Weltkörper dieselbe Oper mit demselben Libretto, denselben Namen des Komponisten, des Textdichters, der Orchestermitglieder und Sänger aufgeführt werden muß.

Die Mächtigkeit des Kontinuums. Wir nannten $\aleph$ die Mächtigkeit der Menge aller reellen Zahlen oder aller Punkte einer Geraden.[1] Nun wird z. B. durch eine Funktion wie $y = \operatorname{tg} x$, wo x das Intervall $-\frac{\pi}{2} < x < \frac{\pi}{2}$, y alle reellen Zahlen durchläuft, die ganze Gerade auf eine endliche Strecke (ohne Endpunkte) umkehrbar eindeutig abgebildet, und eine leichte Modifikation zeigt, daß die Länge dieser Strecke keine Rolle spielt. Jede noch so kleine Strecke ist also, als Punktmenge, ebenfalls von der Mächtigkeit $\aleph$, und jede Punktmenge auf der geraden Linie (lineare Punktmenge), die eine Strecke enthält, hat eine Mächtigkeit, die zugleich $\leqq \aleph$ und $\geqq \aleph$, also $= \aleph$ ist.

Bildet man die Summe einer endlichen Menge oder Folge[2] paarweiser fremder Strecken, so erhält man demgemäß

$$2\aleph = \aleph + \aleph = \aleph,$$

$$n\aleph = \aleph + \aleph + \ldots + \aleph = \aleph,$$

$$\aleph_0 \aleph = \aleph + \aleph + \aleph + \ldots = \aleph.$$

Wir hatten gesehen, daß die Menge der Dezimalbrüche

$$0, a_1 a_2 a_3 \ldots$$

die Mächtigkeit $10^{\aleph_0}$ hat. Nun entspricht jedem solchen Dezimalbruch eine reelle Zahl x des Intervalls $0 \leqq x \leqq 1$, umgekehrt jeder

[1] Wir dürfen jetzt wohl, an die Gewohnheiten der analytischen Geometrie anknüpfend, reelle Zahlen x und Punkte einer Geraden, ebenso geordnete Zahlenpaare (x, y) und Punkte einer Ebene usw. als Synonyma behandeln, obwohl es sich strenggenommen nicht um Identität, sondern nur um Äquivalenz handelt.

[2] z. B. aller Strecken $n-1 < x \leqq n$ $(n = 1, 2, 3, \ldots)$, die zusammen die Halbgerade $x > 0$ ergeben.

solchen reellen Zahl ein Dezimalbruch oder zwei (z. B. $\frac{1}{2} = 0{,}4999\ldots$ oder $= 0{,}5000\ldots$). Ist $\aleph'$ die Mächtigkeit der Menge derjenigen reellen Zahlen, denen zwei Dezimalbrüche entsprechen (übrigens ist $\aleph' = \aleph_0$), so ist also

$$10^{\aleph_0} = \aleph + \aleph'$$

und dies ist einerseits $\geqq \aleph$, andererseits $\leqq \aleph + \aleph = \aleph$, also

$$10^{\aleph_0} = \aleph.$$

Da die Tatsache, daß wir zehn Finger haben, offenbar auf die Mengenlehre ohne Einfluß ist, so können wir statt dekadischer Brüche auch dyadische, triadische usw. zur Darstellung der reellen Zahlen verwenden und erhalten also

(1) $$\aleph = 2^{\aleph_0} = 3^{\aleph_0} = \ldots = (n+1)^{\aleph_0}$$

für jede natürliche Zahl n (während $1^{\aleph_0} = 1$ ist).

Diese Formel (1), die einen merkwürdigen Zusammenhang zwischen $\aleph_0$ und $\aleph$, zwischen der ersten Mächtigkeit und der des Kontinuums enthüllt, ist für die Mengenlehre fundamental und verdient eingehende Betrachtung.

Wir wollen die Darstellung der reellen Zahlen durch dyadische Brüche, die der Formel $\aleph = 2^{\aleph_0}$ entspricht, noch in eine etwas andere Form bringen. Ein dyadischer Bruch

$$x = \frac{d_1}{2} + \frac{d_2}{4} + \frac{d_3}{8} + \ldots, \qquad d_i = 0 \text{ oder } 1$$

stellt eine reelle Zahl x dar ($0 \leqq x \leqq 1$); umgekehrt gehört zu jeder solchen Zahl x ein dyadischer Bruch oder zwei, z. B.

$$\frac{1}{2} = \frac{1}{2} + \frac{0}{4} + \frac{0}{8} + \frac{0}{16} + \ldots = \frac{0}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots,$$

genau wie das bei den dekadischen Brüchen der Fall war. Um diese Zweideutigkeit zu beseitigen, vereinbaren wir, nur solche dyadische Brüche zuzulassen, die *unendlich viele Einsen* enthalten; wonach wir uns also in der letzten Formel für die zweite Darstellung zu entscheiden haben. Mit dieser Verabredung liefert jetzt jeder dyadische Bruch eine reelle Zahl x des Intervalls

$$0 < x \leqq 1$$

und umgekehrt jede solche Zahl x einen dyadischen Bruch.

Schreiben wir dann in der Formel

$$x = \sum_n \frac{d_n}{2^n}$$

nur die Glieder hin, wo $d_n = 1$, und lassen die übrigen mit $d_n = 0$ weg, so erhalten wir

$$x = \left(\frac{1}{2}\right)^{p_1} + \left(\frac{1}{2}\right)^{p_2} + \left(\frac{1}{2}\right)^{p_3} + \ldots,$$

wobei $p_1 < p_2 < p_3 < \dots$ eine Folge wachsender natürlicher Zahlen bedeuten. Endlich formen wir, indem wir

$$a_1 = p_1, \; a_2 = p_2 - p_1, \; a_3 = p_3 - p_2, \dots$$

setzen, dies in

$$(2) \qquad x = \left(\frac{1}{2}\right)^{a_1} + \left(\frac{1}{2}\right)^{a_1 + a_2} + \left(\frac{1}{2}\right)^{a_1 + a_2 + a_3} + \dots$$

um, wobei die $a_1, a_2, a_3, \dots$ eine Folge natürlicher Zahlen bilden. Da wir von dieser Darstellung öfter Gebrauch zu machen haben, wollen wir sie etwas kürzer mit

$$(2) \qquad x = [a_1, a_2, a_3, \dots]$$

bezeichnen. Jetzt entspricht also jeder reellen Zahl x zwischen 0 (exkl.) und 1 (inkl.) umkehrbar eindeutig eine Folge natürlicher Zahlen $a_1, a_2, a_3, \dots$, und auf Grund der bekannten Mächtigkeit der Menge dieser Folgen erhalten wir die Gleichung

$$\aleph = \aleph_0^{\aleph_0}.$$

Zunächst folgt nun aus (1) nach dem Satze § 4, II

$$\aleph = 2^{\aleph_0} > \aleph_0,$$

das Kontinuum (die Menge der reellen Zahlen) ist nicht abzählbar. Die reellen Zahlen lassen sich also nicht in eine Folge bringen. Um uns diesen wichtigen Sachverhalt noch einmal klar zu machen, wobei wir natürlich nur den Grundgedanken des Beweises von § 4, II variieren, zeigen wir, daß es zu jeder Folge reeller Zahlen noch eine weitere, von allen verschiedene reelle Zahl gibt. Sei, etwa in Dezimalbruchform,

$$a = 0, a_1 \, a_2 \, a_3 \dots$$
$$b = 0, b_1 \, b_2 \, b_3 \dots$$
$$c = 0, c_1 \, c_2 \, c_3 \dots$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

eine Folge reeller Zahlen des Intervalls 0, 1. Wir können dann ohne weiteres einen von allen diesen verschiedenen Dezimalbruch

$$x = 0, x_1 \, x_2 \, x_3 \dots$$

angeben, indem wir nur einfach

$$x_1 \neq a_1, \; x_2 \neq b_2, \; x_3 \neq c_3, \dots$$

wählen, und können durch Vermeidung der Ziffern 0 und 9 nebenbei dafür sorgen, daß auch dem Werte nach x von $a, b, c, \dots$ verschieden ausfällt.

Nach den Bemerkungen im Anschluß an § 4, I bleibt, wenn man von den reellen Zahlen endlich oder abzählbar viele entfernt, immer noch eine Menge der Mächtigkeit $\aleph$ zurück. Entfernt man die abzählbare Menge der rationalen Zahlen, so bleibt die Menge der irrationalen, entfernt man die abzählbare Menge

der (reellen) algebraischen Zahlen, so bleibt die Menge der transzendenten Zahlen; beide Mengen haben die Mächtigkeit $\aleph$, und es gibt also, sozusagen, unendlich viel mehr irrationale als rationale, unendlich viel mehr transzendente als algebraische Zahlen.

Wir können auch eine einfache Zuordnung zwischen der Menge aller reellen Zahlen x des Intervalls $0 < x \leqq 1$ und der Menge aller irrationalen Zahlen y desselben Intervalls $0 < y < 1$ angeben, indem wir nämlich der Folge natürlicher Zahlen $(a_1, a_2, a_3, \ldots)$ einerseits nach (2) die Zahl x, andererseits den Kettenbruch

$$y = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

zuordnen.

Wir sahen, daß $\aleph > \aleph_0$. Schon vor vielen Jahren hat G. Cantor die Vermutung ausgesprochen, daß $\aleph$ die auf $\aleph_0$ nächstfolgende Mächtigkeit sei, d. h. daß jede unabzählbare Menge eine Mächtigkeit $\geqq \aleph$ habe. Wir werden später eine solche auf $\aleph_0$ unmittelbar folgende Mächtigkeit, die zweite Mächtigkeit $\aleph_1$, auf anderem Wege kennen lernen. Für diese wissen wir dann, daß $\aleph \geqq \aleph_1$; die Cantorsche Vermutung lautet also, daß $\aleph = \aleph_1$. Der Beweis oder die Widerlegung dieser Vermutung, das Kontinuumproblem, hat bis jetzt allen Bemühungen widerstanden.

Da sich die Punkte der Ebene umkehrbar eindeutig den geordneten Paaren (x, y) reeller Zahlen zuordnen lassen, so hat ihre Menge die Mächtigkeit $\aleph\aleph = \aleph^2$. In ähnlicher elementarer Weise, wie man die Gerade auf eine Strecke (ohne Endpunkte) eineindeutig abbilden konnte, kann man auch die Ebene auf ein Quadrat, ein Rechteck, einen Kreis (d. h. auf das Innere dieser Flächen) abbilden, und erkennt somit, daß jede ebene Punktmenge, die ein Quadrat enthält, jede zweidimensionale Punktmenge (wie wir der Kürze halber sagen wollen, vorbehaltlich späterer eingehender Diskussion des Begriffs Dimension) dieselbe Mächtigkeit $\aleph^2$ hat. Nun ist aber nach (1) und den Potenzgesetzen

$$\aleph^2 = \aleph\aleph = 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0 + \aleph_0} = 2^{\aleph_0} = \aleph,$$

die Ebene hat also nur dieselbe Mächtigkeit wie die Gerade, das Quadrat dieselbe wie die Strecke, eine zweidimensionale Punktmenge dieselbe wie eine eindimensionale. Diese ebenfalls höchst paradoxe Entdeckung G. Cantors, die den Dimensionenbegriff vollkommen umzustoßen und den Unterschied zwischen Linien und Flächen aufzuheben scheint, wollen wir uns auch noch einmal direkt

klar machen, wobei wir ja weiter nichts tun, als die obige abstrakte Formel in die entsprechenden Verknüpfungen zwischen Mengen zurückübersetzen.

Benutzen wir die Darstellung (2), so können wir einem geordneten Paar $p=(x,y)$ von zwei reellen Zahlen

$$x=[a_1, a_2, a_3, \ldots]$$
$$y=[b_1, b_2, b_3, \ldots]$$

unmittelbar eine dritte Zahl

$$t=[a_1, b_1, a_2, b_2, a_3, b_3, \ldots]$$

zuordnen, die hierdurch (wegen der Eindeutigkeit der Darstellung (2), d. h. der umkehrbar eindeutigen Beziehung zwischen x und der Zahlenfolge $a_1, a_2, \ldots$) eine eindeutige Funktion

$$t=F(p)=f(x,y)$$

von p oder x, y wird. Ist umgekehrt

$$t=[c_1, c_2, c_3, c_4, \ldots]$$

gegeben, so erhält man daraus

$$x=[c_1, c_3, c_5, \ldots]$$
$$y=[c_2, c_4, c_6, \ldots]$$

und es werden x, y, p eindeutige Funktionen von t:

$$p=\Phi(t), \quad x=\varphi(t), \quad y=\psi(t).$$

Damit ist also eine umkehrbar eindeutige Beziehung zwischen der Strecke

$$0<t\leqq 1$$

und dem (halbberandeten) Quadrat

$$0<x\leqq 1, \quad 0<y\leqq 1$$

hergestellt. Der Übergang zwischen x, y und t ist ja nichts anderes als der zwischen einem Paar von Folgen und einer einfachen Folge, also die konkrete Deutung der Formel $\aleph_0+\aleph_0=\aleph_0$, auf der eben die Gleichung $\aleph^2=\aleph$ beruhte. Die hier auftretenden „Funktionen" haben allerdings äußerst geringe Ähnlichkeit mit denen, die der Leser aus den Elementen der Analysis kennt und die ja auch als Punktmenge

$$x=\varphi(t), \quad y=\psi(t)$$

eine „Kurve", nicht wie hier eine Fläche liefern würden; von Stetigkeit, geschweige Differenzierbarkeit ist keine Rede, und bei dem Versuche, sich ihren Verlauf vorzustellen, versagt die Anschauung.

Wir werden in der Lehre von den Funktionen auf diese Dinge zurückzukommen haben.

Daß auch der Raum oder die Menge der Zahlenkomplexe (x, y, z), allgemein der n-dimensionale Raum oder die Menge der Komplexe $(x_1, x_2, \ldots, x_n)$ reeller Zahlen nur die Mächtigkeit $\aleph$ hat, ist ja nun gemäß der Formel

$$\aleph^n = \aleph\,\aleph \ldots \aleph = \aleph$$

nichts Neues. Aber sogar die Menge der Folgen reeller Zahlen

$$(x_1, x_2, x_3, \ldots)$$

oder, wenn man so sagen will, der Punkte des $\aleph_0$-dimensionalen Raumes hat keine höhere Mächtigkeit, denn es ist ja

$$\aleph^{\aleph_0} = \left(2^{\aleph_0}\right)^{\aleph_0} = 2^{\aleph_0 \aleph_0} = 2^{\aleph_0} = \aleph.$$

Wollen wir auch dies durch eine umkehrbar eindeutige Zuordnung zwischen der Menge jener Folgen und der Menge der reellen Zahlen t verifizieren, so nehmen wir wieder, unter Beschränkung auf die Intervalle

$$0 < x_n \leqq 1, \quad 0 < t \leqq 1,$$

die Darstellung (2) zu Hilfe. Aus

$$\begin{aligned} x_1 &= [a_{11}, a_{12}, a_{13}, \ldots] \\ x_2 &= [a_{21}, a_{22}, a_{23}, \ldots] \\ x_3 &= [a_{31}, a_{32}, a_{33}, \ldots] \\ &\ldots\ldots\ldots \end{aligned}$$

erhalten wir dann nach dem früheren Diagonalverfahren (S. 18), das eine Doppelfolge in eine einfache verwandelt:

$$\begin{aligned} t &= [a_{11}, a_{12}, a_{21}, a_{13}, a_{22}, a_{31}, \ldots] \\ &= f(x_1, x_2, x_3, \ldots) \end{aligned}$$

als eindeutige Funktion der Folge $(x_1, x_2, x_3, \ldots)$, und umgekehrt aus

$$t = [c_1, c_2, c_3, c_4, \ldots]:$$

$$\begin{aligned} x_1 &= [c_1, c_2, c_4, \ \ldots] = \varphi_1(t) \\ x_2 &= [c_3, c_5, c_8, \ \ldots] = \varphi_2(t) \\ x_3 &= [c_6, c_9, c_{13}, \ldots] = \varphi_3(t) \\ &\ldots\ldots\ldots\ldots \end{aligned}$$

als eindeutige Funktionen von t.

Bilden $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3, \ldots$ eine Folge von Mächtigkeiten, die zwischen 2 und $\aleph$ liegen $(2 \leqq \mathfrak{a}_n \leqq \aleph)$, so ist

$$\mathfrak{P}\,\mathfrak{a}_n = \mathfrak{a}_1\, \mathfrak{a}_2\, \mathfrak{a}_3 \ldots = \aleph;$$

denn dieses Produkt ist zugleich $\geqq 2^{\aleph_0} = \aleph$ und $\leqq \aleph^{\aleph_0} = \aleph$. Die Formel gilt auch noch, wenn unendlich viele $\mathfrak{a}_n$ zwischen 2 und $\aleph$ liegen und die übrigen $= 1$ sind. Z. B. ist

$$1 \cdot 2 \cdot 3 \cdot 4 \ldots = \aleph.$$

Wir konstatieren noch, daß $\aleph$ sicher nicht die Summe einer aufsteigenden Folge von Mächtigkeiten sein kann, denn für eine solche Summe $\mathfrak{a}$ gilt nach § 4, IV $\mathfrak{a} < \mathfrak{a}^{\aleph_0}$, während $\aleph = \aleph^{\aleph_0}$ ist.

Die Mächtigkeit $2^{\aleph}$. Sie ist wieder $> \aleph$. Nach der Definition ist sie die Mächtigkeit der Menge aller eindeutigen Funktionen $f(x)$ einer reellen Variablen x, wenn diese Funktion zwei Werte, z. B. 1 und 2, annehmen kann. Sie ist auch die Mächtigkeit des Systems aller Teilmengen einer Menge von der Mächtigkeit $\aleph$, z. B. des Systems aller linearen oder ebenen oder räumlichen Punktmengen.

Die Menge aller reellen Funktionen einer reellen Variablen, wo also $f(x)$ alle reellen Zahlen als Werte annehmen kann, hat auch nur die Mächtigkeit

$$\aleph^{\aleph} = (2^{\aleph_0})^{\aleph} = 2^{\aleph_0 \aleph} = 2^{\aleph}.$$

Auch die Menge von Funktionensystemen

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_m) \\ y_2 &= f_2(x_1, x_2, \ldots, x_m) \\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ y_n &= f_n(x_1, x_2, \ldots, x_m) \end{aligned}$$

hat nur diese Mächtigkeit. Denn man fasse die Komplexe

$$x = (x_1, x_2, \ldots, x_m), \quad y = (y_1, y_2, \ldots, y_n)$$

als Punkte eines m- resp. n-dimensionalen Raumes auf; die Menge der Funktionen $y = f(x)$ hat die Mächtigkeit $\aleph^{\aleph} = 2^{\aleph}$. Dasselbe gilt auch noch für abzählbar viele reelle Funktionen abzählbar vieler reeller Variabler.

Spezielle Funktionenklassen können natürlich geringere Mächtigkeit haben. So hat, wie wir hier vorwegnehmen, die Menge der stetigen Funktionen $f(x)$ nur die Mächtigkeit $\aleph$. Denn eine stetige Funktion ist durch ihre Werte $f(r)$ an den rationalen Stellen r bestimmt. Da die Menge aller reellen Funktionen $f(r)$ (auch die mitgerechnet, die keine stetige Funktion $f(x)$ liefern) die Mächtigkeit $\aleph^{\aleph_0} = \aleph$ hat, so hat die Menge aller stetigen Funktionen eine Mächtigkeit $\leqq \aleph$, andererseits auch $\geqq \aleph$, da schon die stetigen Funktionen $f(x) =$ constans eine Menge von der Mächtigkeit $\aleph$ bilden, also schließlich die Mächtigkeit $\aleph$.

Wir schließen dies Kapitel mit einer Formelzusammenstellung, in der man alle Summen $\mathfrak{a} + \mathfrak{b}$, Produkte $\mathfrak{a}\mathfrak{b}$ und Potenzen $\mathfrak{a}^{\mathfrak{b}}$ beisammen hat, die sich aus einer natürlichen Zahl n, aus $\aleph_0$ und $\aleph$ bilden lassen, mit Ausnahme der Kombinationen aus zwei natürlichen Zahlen.

(α) $$1^{\aleph_0} = 1^{\aleph} = 1.$$

(β) $$n + \aleph_0 = n \cdot \aleph_0 = \aleph_0^n = \aleph_0 + \aleph_0 = \aleph_0 \cdot \aleph_0 = \aleph_0.$$

(γ) $$\begin{aligned}(n+1)^{\aleph_0} = \aleph_0^{\aleph_0} &= n + \aleph = n \cdot \aleph = \aleph^n \\ &= \aleph_0 + \aleph = \aleph_0 \cdot \aleph = \aleph^{\aleph_0} \\ &= \aleph + \aleph = \aleph \cdot \aleph = \aleph.\end{aligned}$$

(δ) $$(n+1)^{\aleph} = \aleph_0^{\aleph} = \aleph^{\aleph} = 2^{\aleph}.$$

Soweit diese Formeln noch nicht aufgetreten sind, sind sie mit Anwendung des Äquivalenzsatzes leicht zu beweisen.

Viertes Kapitel.

Geordnete Mengen. Ordnungstypen.

§ 1. Ordnung.

Unser Ziel, auch unendliche Mengen zu zählen, haben wir bisher nur sehr unvollkommen erreicht. Wir haben zwar jeder Menge eine Mächtigkeit oder Kardinalzahl zugewiesen, aber sind noch nicht sicher, daß zwei verschiedene Mächtigkeiten stets miteinander vergleichbar sind, daß die eine die größere und die andere die kleinere ist. Ferner ist das Zählen einer endlichen Menge noch mehr als die bloße Feststellung des Endresultats, daß diese Menge aus so und so vielen Dingen bestehe: während des Zählaktes ordnen wir ja den Elementen selbst Zahlzeichen oder Zahlworte zu. Hieran knüpft sich die doppelte Bedeutung der Zahl als Ordinalzahl oder Nummer eines Elements und als Kardinalzahl oder Elementenvorrat einer Menge; während wir zählen 1, 2, 3, 4, 5, benennen wir die einzelnen Gegenstände, um dann im Ergebnis zu konstatieren, daß die Menge aus 5 Gegenständen besteht. In beiden Beziehungen fehlt uns für die unendlichen Mengen noch ein Instrument, wie es für die endlichen die natürliche Zahlenreihe ist: ein System von Zeichen in festgelegter Reihenfolge, Zeichen, mit denen wir die Elemente einer Menge benennen und deren Anordnung nicht nur in einer Menge das frühere Element von dem späteren, sondern auch bei verschiedenen Mengen die kleinere von der größeren zu unterscheiden gestattet.

Wir werden bei dem Versuch, diese Bemerkungen über den Zählprozeß für unendliche Mengen zu verwerten, schrittweise vorgehen müssen und in erster Linie konstatieren, daß Zählen ein

Ordnen ist, ein Vorgang, durch den von zwei verschiedenen Elementen das eine vor das andere, dieses hinter jenes gereiht wird.

Dieser Prozeß der Ordnung läßt sich nun unmittelbar auf beliebige Mengen übertragen. Wir ordnen eine Menge, indem wir eine Vorschrift geben, nach der von zwei verschiedenen ihrer Elemente, a und b, das eine als das frühere, das andere als das spätere charakterisiert wird. Soll a das frühere, b das spätere sein, so schreiben wir

$$a < b \quad \text{oder} \quad b > a;$$

wir lesen dies „a vor b", „b nach a", finden es aber entbehrlich, hierfür andere Zeichen als die gewöhnlichen für kleiner und größer zu benutzen. Ist umgekehrt b das frühere, a das spätere Element, so ist zu schreiben

$$b < a \quad \text{oder} \quad a > b.$$

Hierbei ist aber noch eine wichtige Bestimmung zu beachten: wenn a vor b und b vor c steht, so soll auch a vor c stehen, also

$$\text{aus } a < b,\ b < c \text{ folgt } a < c.$$

Das Zeichen $<$ (und ebenso $>$) soll, wie man sagt, das transitive Gesetz befolgen.

Wie kann man eine solche ordnende Vorschrift geben? Nun, die Sache ist uns gar nichts Neues, sondern subsumiert sich unter unseren allgemeinen Funktionenbegriff. Wir haben auf der einen Seite die Menge der geordneten Paare

$$p = (a, b),$$

die aus zwei verschiedenen Elementen ($a \neq b$) einer Menge gebildet sind und wobei also p von dem umgekehrten Paar

$$p^* = (b, a)$$

verschieden ist; auf der anderen Seite haben wir eine Menge, aus zwei Elementen bestehend, die diesmal nur in anscheinend sonderbarer Weise mit $<$ und $>$ bezeichnet sind. Nun bilden wir eine eindeutige Funktion $f(p)$, die einen der „Werte" $<$ oder $>$ haben soll, nur daß wir kürzer

$$\text{statt } f(p) = < \ : \ a < b,$$
$$\text{statt } f(p) = > \ : \ a > b$$

schreiben, und daß $f(p)$ nicht völlig willkürlich ist, sondern $f(p^*)$ immer den anderen Wert als $f(p)$ haben soll, also $f(p^*) \neq f(p)$, und überdies das transitive Gesetz gelten soll: wenn $p = (a, b)$, $q = (b, c)$, $r = (a, c)$ und $f(p) = f(q)$, so ist auch $f(p) = f(r)$.

Noch übersichtlicher wird dies, wenn wir uns erinnern, daß eine zweiwertige Funktion eine Zerlegung des Arguments in zwei komplementäre Teilmengen bestimmt und umgekehrt. Wir haben

dann einfach folgendes: es sei eine Menge P geordneter Paare und ihr Komplement P^* gegeben, derart, daß $P+P^*$ die Menge aller geordneten Paare (aus verschiedenen Elementen der Menge A) ist. Diese Mengen sollen folgenden speziellen Charakter haben:

(α) von zwei inversen Paaren p, p^* gehört das eine zu P und das andere zu P^*,

(β) gehört $p=(a,b)$ und $q=(b,c)$ zu P, so auch $r=(a,c)$.

Drückt man dann die Tatsache, daß $p=(a,b)$ zu P resp. zu P^* gehört, durch

$$a < b \quad \text{resp.} \quad a > b$$

aus, so bestimmt P eine Ordnung der Menge A, wie umgekehrt jede Ordnung die Menge P derjenigen Paare (a, b), für die $a<b$ ist, damit auch die andere Menge P^* bestimmt und die Eigenschaften (α), (β) realisiert. P heiße die ordnende Paarmenge für die geordnete Menge A. Um ein ganz einfaches Beispiel zu geben: wenn wir aus den Paaren der Menge $\{a, b, c\}$ die folgenden Mengen bilden

$$P\ : (a, b), (b, c), (a, c)$$
$$P^*: (b, a), (c, b), (c, a),$$

so ist unseren Forderungen (α), (β) genügt und die Ordnung $a<b<c$ definiert.

Wir bezeichnen in diesem und den beiden folgenden Kapiteln geordnete Mengen mit großen lateinischen Buchstaben; unter der Gleichung $A=A'$ verstehen wir dann nicht bloß, wie früher, daß beide Mengen dieselben Elemente haben, sondern auch dieselbe Ordnung, d. h. daß die ordnenden Paarmengen P, P' identisch sind.

Wie man sieht, erfüllt die Menge P^* genau dieselben Bedingungen wie P. Nimmt man sie statt P als ordnende Paarmenge, so erhält man die zu A invers geordnete Menge A^*; ist nämlich in A $a<b$, so ist in A^* $a>b$.

Die Ordnung einer Menge hat mit Zeit, Größe usw. nichts zu schaffen, wenn wir auch die zeitlich oder räumlich klingenden Präpositionen „vor" und „nach" verwenden, sondern kann willkürlich festgesetzt werden. Auch Mengen, die wir uns gewöhnlich in einer festen, „natürlichen" Ordnung denken, können nach Belieben umgeordnet werden. Wir geben einige Beispiele für Ordnung der Menge der natürlichen Zahlen, wobei wir der Kürze wegen (wie wir dies öfter tun werden) die Ordnung nur durch die Reihenfolge des Schreibens andeuten, so daß, für $a<b$, a links von b steht; die beigesetzten Zeichen ω usw. werden später ihre Erklärung finden:

$$(\omega)\quad 1\ 2\ 3\ 4 \cdots$$
$$(\omega^*)\quad \cdots 4\ 3\ 2\ 1$$

$(\omega+\omega)$ 1 3 5 7 . . . 2 4 6 8 . . .

$(\omega+\omega^*)$ 1 3 5 7 . . . . . . 8 6 4 2

$(\omega^*+\omega)$ . . . 8 6 4 2 1 3 5 7 . . .

$(\omega\omega)$ 1 2 4 . . . 3 5 8 . . . 6 9 13 . . . .

Zuerst stehen die Zahlen in „natürlicher“ Ordnung, dann umgekehrt; in der dritten Zeile die ungeraden vor den geraden, beide in natürlicher Ordnung; in der vierten und fünften stehen die ungeraden Zahlen in natürlicher, die geraden in umgekehrter Ordnung, das eine Mal die ungeraden vor den geraden, das andere Mal umgekehrt. Die letzte Zeile ist entstanden, indem man die Zahlenfolge nach bekanntem Diagonalverfahren in eine Doppelfolge

1	2	4	7	11	.	.	.
3	5	8	12	.	.	.	.
6	9	13	.	.	.	.	.
10	14	.	.	.	.	.	.
15	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.

bringt und die Elemente einer Zeile von links nach rechts, die ganzen Zeilen von oben nach unten ordnet.[1]

Wir wollen noch auf eine dritte, formal wieder etwas andere Definition der Ordnung verweisen. Die Menge P der Paare bestimmt (Kap. II, § 4) die Menge $F(a)$ derjenigen Elemente b, die als zweite Elemente mit dem ersten Element a verbunden auftreten; oder die Menge $F(a)$ derjenigen Elemente von A, die $> a$. Umgekehrt ist durch Zuordnung einer Teilmenge $F(a)$ von A zu jedem Element a unter geeigneten Bedingungen eine Ordnung definiert, und zwar genügt es folgende Vorschrift zu geben:

Wenn b Element von $F(a)$ ist, so ist $F(b)$ echte Teilmenge von $F(a)$; für $a \neq b$ ist entweder b Element von $F(a)$ oder a Element von $F(b)$.

Drücken wir dann die Tatsache, daß b Element von $F(a)$ ist, durch $a < b$ aus, so folgt: für $a \neq b$ ist entweder $a < b$ und $F(b) \subset F(a)$, oder $b < a$ und $F(a) \subset F(b)$, also beide Fälle schließen einander aus. Für $a < b$, $b < c$ ist c Element von $F(b) \subset F(a)$, also auch c Element von $F(a)$ und daher $a < c$. Das System der Mengen $F(a)$ definiert also eine Ordnung der Menge A.

Zwei geordnete Mengen A, B heißen gleichgeordnet oder ähnlich, in Zeichen

$$A \simeq B \quad \text{oder} \quad B \simeq A,$$

---

[1] Ebenso könnte man z. B. die reellen Zahlen statt in natürlicher Reihenfolge auch so ordnen, daß jede rationale Zahl vor jeder irrationalen steht, während die rationalen unter sich und die irrationalen unter sich in natürlicher Ordnung stehen.

wenn es eine umkehrbar eindeutige Beziehung $b = f(a)$, $a = \varphi(b)$ zwischen ihnen gibt, bei der die Ordnung entsprechender Elemente dieselbe, d. h. mit $a < a'$ zugleich $f(a) < f(a')$ ist.

Z. B. sind die beiden in natürlicher Ordnung genommenen Zahlenfolgen

$$1\ 2\ 3\ 4\ \ldots \quad \text{und} \quad 2\ 3\ 4\ 5\ \ldots$$

ähnlich, da man jeder Zahl a der ersten die Zahl $b = a + 1 = f(a)$ der zweiten zuordnen kann.

Ist $A \simeq B$, $B \simeq C$, so ist auch $A \simeq C$.

Das den ähnlichen Mengen Gemeinsame bezeichnen wir als Ordnungstypus, wie wir das den äquivalenten Mengen Gemeinsame als Mächtigkeit bezeichneten. Wir ordnen nämlich jeder Menge A ein Zeichen α zu, derart, daß ähnlichen Mengen und nur solchen dasselbe Zeichen entspricht, daß also mit $A \simeq B$ zugleich $\alpha = \beta$ und umgekehrt mit $\alpha = \beta$ zugleich $A \simeq B$ ist. Dieses Zeichen α heißt der Ordnungstypus (oder Typus) der Menge A.

Ähnlichkeit bedingt Äquivalenz, aber nicht umgekehrt; aus $A \simeq B$ folgt $A \sim B$, aus $\alpha = \beta$ folgt $\mathfrak{a} = \mathfrak{b}$. Wir dürfen daher auch sagen, ein Typus α habe die Mächtigkeit $\mathfrak{a}$.

Hat A den Typus α, so soll die invers geordnete Menge A^* den Typus α^* haben.

Eine endliche Menge aus n Elementen ($n > 1$) kann zwar auf verschiedene Weise geordnet werden, indessen sind die sämtlichen so entstehenden geordneten Mengen einander ähnlich. Bei jeder Anordnung muß es nämlich ein erstes Element geben, das vor allen anderen steht; denn in einer Menge ohne erstes Element gibt es zu jedem Element ein früheres, also enthält die Menge eine absteigende Folge $a_1 > a_2 > a_3 > \ldots$ und ist demnach unendlich. Es gibt also in einer endlichen geordneten Menge A ein erstes Element a_1, in $A - \{a_1\}$ wieder ein erstes a_2 usw., d. h. A erscheint in der Anordnung $a_1 < a_2 < \ldots < a_n$ und ist mit der in natürlicher Reihenfolge genommenen Menge der natürlichen Zahlen $1, 2, \ldots, n$ ähnlich. Wir bezeichnen den Typus dieser Zahlenmenge mit n, so daß jede geordnete Menge aus n Elementen den Typus n hat. Daß wir hier Kardinalzahl und Ordnungstypus gleich bezeichnen, ist unbedenklich.[1] — Eine Menge aus einem einzigen Element, die also,

[1] Treten in einer Gleichung unendliche Mengen auf, bei denen ja zwischen Mächtigkeit und Typus unterschieden wird, so sieht man ihr eben dadurch an, ob sie als Gleichung zwischen Typen oder Mächtigkeiten gemeint ist, z. B.

$$1 + 1 + 1 + \ldots = \omega, \quad 1 + 1 + 1 + \ldots = \aleph_0.$$

Treten nur endliche Mengen auf, so ist es gleichgültig, ob man die Zahlen als Typen oder Mächtigkeiten auffaßt.

wie man es nun auffassen will, wegen Mangels an Paaren nicht geordnet werden kann oder eo ipso schon geordnet ist, erhalte den Typus 1, die Nullmenge den Typus 0.

Den Typus der natürlichen Zahlenmenge in natürlicher Ordnung bezeichnen wir mit ω, in umgekehrter Ordnung also mit ω^*.

Im Gegensatz zu den endlichen Mengen kann jede unendliche Menge zu mehreren (unendlich vielen) Mengen von verschiedenen Typen geordnet werden. Die natürliche Zahlenreihe läßt z. B. außer den schon oben angeführten folgende Anordnungen zu, denen wir wieder nachher zu erklärende Typenzeichen beisetzen:

$$\begin{array}{ll} (\ \omega\ ) & 1\ 2\ 3\ 4 \cdots \\ (\omega+1) & 2\ 3\ 4\ 5 \cdots 1 \\ (\omega+2) & 3\ 4\ 5\ 6 \cdots 1\ 2 \\ (\omega+3) & 4\ 5\ 6\ 7 \cdots 1\ 2\ 3 \\ \cdots & \cdots \end{array}$$

Bei der ersten Anordnung gibt es kein Element, dem nicht unendlich viele nachfolgen, bei der zweiten gibt es ein solches (1), bei der dritten zwei (1, 2), bei der vierten drei (1, 2, 3) usw.; diese Typen sind also alle verschieden. Für eine allgemeine unendliche Menge kann man eine ähnliche Betrachtung anstellen.

§ 2. Verknüpfungen geordneter Mengen.

Die geordnete Menge A prägt auch allen ihren Teilmengen A' eine bestimmte Ordnung auf, diejenige nämlich, die die Elemente von A' als Elemente von A haben, oder die man erhält, wenn man aus der ordnenden Paarmenge P die Menge P' derjenigen Paare aussondert, die nur aus Elementen von A' bestehen. In diesem Sinne kann man auch von der geordneten Summe und dem geordneten Durchschnitt

$$S = \mathfrak{S}(A_1, A_2, \ldots),\ D = \mathfrak{D}(A_1, A_2, \ldots)$$

geordneter Mengen A_i sprechen, wenn nämlich alle diese Mengen Teilmengen einer geordneten Menge A sind.

Eine unendliche Menge kann mit einer echten Teilmenge nicht nur äquivalent, sondern auch ähnlich sein. Z. B. ist die Zahlenreihe 1 2 3 4 ... in natürlicher Ordnung mit der Menge der Quadratzahlen 1 4 9 16 ... ähnlich. Die Menge aller reellen Zahlen in natürlicher Ordnung ist der Menge der positiven Zahlen oder der Menge der reellen Zahlen eines Intervalls ohne Endpunkte ähnlich, wie die Funktionen

$$y = \log x \quad (x > 0)$$
$$y = \operatorname{tg} x \quad \left(-\frac{\pi}{2} < x < \frac{\pi}{2}\right)$$

lehren.

Haben wir dagegen zwei geordnete Mengen A, B, die noch nicht als Teilmengen einer umfassenden geordneten Menge erkannt sind, so ist eine Definition der geordneten Mengen

$$S = \mathfrak{S}(A, B) \quad \text{und} \quad D = \mathfrak{D}(A, B)$$

im allgemeinen unmöglich. Nur wenn die Anordnung, die D als Teilmenge von A empfängt, mit der von B empfangenen übereinstimmt, ist D dadurch als geordnete Menge eindeutig definiert; und nur in diesem Falle kann man auch S, im allgemeinen aber auf verschiedene Arten, so ordnen, daß A und B als Teilmengen von S eben die Ordnung empfangen, die sie schon hatten.

Wir wollen uns auf den Fall $D = 0$ beschränken und unter den dann noch möglichen Ordnungen von S eine bestimmte auszeichnen.

Unter der Summe $A + B$ zweier geordneter, fremder Mengen versteht man die Summe der beiden Mengen A, B in derjenigen Ordnung, die man erhält, wenn man jedes Element a vor jedes Element b setzt, aber die Ordnung der Elemente a unter sich und die der Elemente b unter sich bestehen läßt.

Anders ausgedrückt: man vereinigt die ordnenden Paarmengen von A und B und fügt noch die Menge der Paare (a, b) hinzu. Man sieht unmittelbar: wenn $A \simeq A'$, $B \simeq B'$, und wenn A mit B, A' mit B' fremd ist, so ist

$$A + B \simeq A' + B'.$$

Dies berechtigt, die Summe $\alpha + \beta$ zweier Typen als Typus der Summe $A + B$ zu definieren, wenn A, B zwei fremde Mengen vom Typus α, β sind.

Hier gilt das kommutative Gesetz nicht mehr; die geordneten Mengen $B + A$ und $A + B$ sind, obwohl als Mengen schlechthin aus denselben Elementen zusammengesetzt, verschieden und im allgemeinen auch nicht ähnlich ($\beta + \alpha \neq \alpha + \beta$).

Wir schließen sofort die Definition einer Summe mit beliebigem Argument an. Den Elementen i einer geordneten Menge J weisen wir paarweise fremde, geordnete Mengen A_i zu und verstehen dann unter der Summe

$$S = \sum_i^J A_i$$

die Summe der Mengen A_i in folgender Anordnung: zwei Elemente von A_i behalten in S die Ordnung, die sie in A_i hatten; hingegen ist für ein Element a_i von A_i und ein Element a_k von A_k ($i \neq k$)

$$a_i \lesseqgtr a_k \text{ in } S, \text{ je nachdem } i \lesseqgtr k \text{ in } J.$$

Die ordnende Paarmenge von S entsteht also durch Vereinigung

aller ordnenden Paarmengen der A_i und Hinzufügung der Paare (a_i, a_k) für $i < k$.

Hier gilt wieder das assoziative Gesetz, nämlich für

$$J = \sum_m^M J_m$$

ist

$$\sum_i^J A_i = \sum_m^M \sum_i^{J_m} A_i .$$

In der Tat, bezeichnen wir die linke und rechte Seite der letzten Formel mit L und R, so ist die Identität der Elemente beider Mengen schon bekannt. Um die Identität der Ordnung einzusehen. unterscheiden wir für zwei verschiedene Elemente a, b folgende Fälle:

(1) Sie gehören derselben Menge A_i an: ihre Ordnung ist dann in L und R dieselbe wie in A_i.

(2) Sie gehören verschiedenen Mengen A_i an, a zu A_i und b zu A_k. Die Ordnung von a, b in L ist dieselbe wie die von i, k in J. Dieser Fall spaltet sich in zwei Unterfälle:

(21) Die Indices i, k gehören derselben Menge J_m an. Die Ordnung von a, b ist in R dieselbe wie die von i, k in J_m, und diese ist auch die Ordnung von i, k in J.

(22) Die Indices i, k gehören verschiedenen Mengen J_m an, i zu J_m und k zu J_n. Die Ordnung von a, b in R ist dieselbe wie die von m, n in M, also dieselbe wie die Ordnung von i, k in J.

Spezielle Fälle des assoziativen Gesetzes sind

$$A + (B + C) = (A + B) + C = A + B + C.$$

Ein kommutatives Gesetz gilt nicht; schon wenn wir die Zuordnung der A_i zu den i beibehalten und nur aus J eine andere Menge J' mit denselben Elementen in anderer Ordnung bilden, sind die geordneten Mengen

$$\sum_i^J A_i \text{ und } \sum_i^{J'} A_i$$

verschieden und im allgemeinen unähnlich. Nur die Ersetzung von J durch eine ähnliche Menge K ist statthaft: bezeichnet $i = \varphi(k)$ eine umkehrbar eindeutige, ähnliche Abbildung von J auf K, und ordnen wir dem Element k die Menge $B_k = A_{\varphi(k)}$ zu, so ist

$$\sum_k^K B_k = \sum_i^J A_i .$$

Ferner geht die Summe

$$\sum_i^J A_i$$

bei Ersetzung der A_i durch ähnliche, wieder paarweise fremde

Mengen in eine ähnliche Summe über; dies berechtigt, als Typensumme

$$\sum_i^J \alpha_i$$

den Typus der Mengensumme

$$\sum_i^J A_i$$

zu definieren, wo die A_i paarweise fremde Mengen vom Typus α_i sind.

Wenn wir in

$$\sum_i^J A_i$$

einige Indices hervorheben und demgemäß die Summe unter Verwendung von Pluszeichen schreiben wollen, so soll für die Reihenfolge der Indices in J wie die der Summanden die Schreibweise von links nach rechts maßgebend sein; wir würden also, um auszudrücken, daß $i < k < l$, schreiben:

$$J = \{\ldots, i, \ldots, k, \ldots, l, \ldots\},$$

$$\sum_i^J A_i = \ldots + A_i + \ldots + A_k + \ldots + A_l + \ldots$$

Die Punkte deuten die eventuelle Anwesenheit weiterer Elemente vor i, zwischen i und k usw. an; das Fehlen solcher Punkte würde also bezeichnen, daß etwa i das erste, k das auf i unmittelbar folgende Element von J ist usw.

Beispiele. ω war der Typus der natürlichen Zahlenmenge in natürlicher Ordnung. Setzen wir dieser ein Element 0 voran, so entsteht der Typus $1 + \omega$ der Menge $\{0, 1, 2, \ldots\}$, die aber der natürlichen Zahlenreihe ähnlich ist. Also haben wir

$$1 + \omega = \omega$$

und ebenso, etwa nach dem assoziativen Gesetz,

$$2 + \omega = (1 + 1) + \omega = 1 + (1 + \omega) = 1 + \omega = \omega,$$

$$n + \omega = \omega.$$

Setzen wir dagegen der Zahlenreihe ein Element nach, so entsteht der Typus $\omega + 1$ der Menge $\{1, 2, 3, \ldots, 0\}$, die ein letztes Element hat und der Zahlenreihe demnach nicht ähnlich ist, also ist $\omega + 1 \neq \omega$ und

$$\omega + 1 \neq 1 + \omega,$$

was die Ungültigkeit des kommutativen Gesetzes beleuchtet. Die geordneten Mengen auf S. 74 von den Typen $\omega, \omega + 1, \omega + 2, \omega + 3, \ldots$ sind sämtlich verschieden. Auf S. 71, 72 haben wir Mengen der Typen $\omega, \omega^*, \omega + \omega, \omega + \omega^*, \omega^* + \omega$ angegeben, die, wie man sich leicht klar macht, ebenfalls untereinander und von den Typen $\omega + n$

verschieden sind. $\omega^* + \omega$ ist der Typus der Reihe der ganzen Zahlen

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

in natürlicher Ordnung.

Ordnen wir jeder natürlichen Zahl i eine natürliche Zahl α_i zu, so ist

$$\sum_i \alpha_i = \alpha_1 + \alpha_2 + \alpha_3 + \ldots = \omega;$$

denn wir können als Mengen A_i die folgenden Mengen natürlicher Zahlen wählen:

$$A_1 = \{1, \ldots, \alpha_1\}$$
$$A_2 = \{\alpha_1 + 1, \ldots, \alpha_1 + \alpha_2\}$$
$$A_3 = \{\alpha_1 + \alpha_2 + 1, \ldots, \alpha_1 + \alpha_2 + \alpha_3\}$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

wonach $\sum_i A_i$ einfach die natürliche Zahlenreihe wird. Z. B. ist

$$1 + 1 + 1 + \ldots = \omega,$$
$$2 + 2 + 2 + \ldots = \omega,$$
$$1 + 2 + 3 + \ldots = \omega.$$

Produkte endlich vieler Faktoren. Als das Produkt zweier Mengen A, B hatten wir im wesentlichen[1] die Menge der geordneten Paare (a, b) definiert, wo a die Menge A, b die Menge B durchläuft. Um im Falle geordneter Mengen auch das Produkt zu ordnen, bietet sich ein Verfahren dar, das man mit glücklichem Ausdruck die lexikographische Anordnung genannt hat. Die Reihenfolge der Worte in einem Lexikon ist ja die, daß man die Worte zunächst nach der alphabetischen Reihenfolge der ersten Buchstaben ordnet, bei Übereinstimmung dieser nach der Reihenfolge der zweiten Buchstaben, bei Gleichheit auch dieser nach der Reihenfolge der dritten Buchstaben usw. Übertragen wir dies auf unsere Paare (a, b), worin a das erste, b das zweite Element ist, so werden wir also definieren:

$$(a, b) < (a', b'),$$

wenn entweder

$$a < a' \text{ in } A$$

oder

$$a = a', \; b < b' \text{ in } B.$$

Das so geordnete Produkt bezeichnen wir mit BA; es ist von der Menge AB der lexikographisch geordneten Paare (b, a) zu unterscheiden, mit der es äquivalent, aber im allgemeinen nicht ähnlich

---

[1] Eigentlich war eine Zuordnung zu einem zweigliedrigen Argument, etwa $\{1, 2\}$, zu verabreden: wir dachten uns diese durch geordnete Schreibweise ausgedrückt.

ist.[1] Auch bei der Multiplikation geordneter Mengen gilt also kein kommutatives Gesetz. Ersetzen wir A, B durch ähnliche Mengen A', B', so geht auch das Produkt BA in ein ähnliches $B'A'$ über. Dies berechtigt, als Produkt $\beta\alpha$ zweier Typen den Typus des Produkts BA zu definieren, wenn A, B irgend zwei Mengen von den Typen α, β sind.

Ist P irgend eine Menge geordneter Paare (a, b), wobei a eine Menge A und, bei festem a, b eine Menge B_a durchläuft, so können wir, falls A und jedes B_a geordnet ist, auch die Menge P lexikographisch ordnen.[2] Ist P_a die Menge der Paare mit festem a, so ist offenbar

$$P = \sum_a^A P_a, \quad P_a \simeq B_a.$$

Wenn alle $B_a = B$ sind, so ist $P = BA$; bezeichnet man die Typen von A, B mit α, β, so ist also

$$\sum_a^A \beta = \beta\alpha,$$

die Summe gleicher Summanden β, über ein Argument vom Typus α, gibt das Produkt $\beta\alpha$. Diese Beziehung zwischen Summe und Produkt ist nützlich, um sich die Bedeutung der Reihenfolge der Faktoren in einem Produkt einzuprägen: $\beta\alpha$ wird erhalten, indem man für jedes Element von A eine Menge vom Typus β setzt, „β in α einsetzt".

Aus dem assoziativen Summengesetz folgt für die oben erklärte Menge P:

$$\sum_a^A P_a = \sum_i^J \sum_a^{A_i} P_a, \text{ falls } A = \sum_i^J A_i$$

und insbesondere für Gleichheit aller $B_a = B$:

$$BA = \sum_i^J BA_i$$

oder

$$B \sum_i^J A_i = \sum_i^J BA_i.$$

---

[1] Die Reihenfolge der Faktoren ist natürlich konventionell; G. Cantor selbst hat die Bezeichnungsweise gewechselt und wir haben uns hier an seine zweite, auch sonst ziemlich allgemein akzeptierte Schreibart angeschlossen. Es wäre übrigens auch möglich, AB als geordnete Menge der Paare (a, b) zu definieren, die aber dann „antilexikographisch" nicht nach dem ersten, sondern nach dem letzten differierenden Element zu ordnen wäre, d. h. so, daß $(a, b) < (a', b')$, wenn entweder $b < b'$ oder $b = b'$, $a < a'$. Diese Modifikation wäre für die spätere Ausdehnung des Produktbegriffs weniger zweckmäßig.

[2] Die verschiedenen B_a brauchen nicht einer und derselben geordneten Menge anzugehören.

Ein Produkt, dessen letzter (zweiter) Faktor eine Summe ist, kann also in eine Summe von Produkten aufgelöst werden: das ist der Teil des distributiven Gesetzes zwischen Summe und Produkt, der bei geordneten Mengen bestehen bleibt. Insbesondere ist

$$B(A_1 + A_2) = BA_1 + BA_2.$$

Die Mengen $(B_1 + B_2)A$ und $B_1 A + B_2 A$ sind dagegen zwar als ungeordnete Mengen identisch, aber als geordnete verschieden und im allgemeinen unähnlich (hierzu wie zur Ungültigkeit des kommutativen Gesetzes vgl. die späteren Beispiele).

Analog definieren wir das Produkt aus einer endlichen Faktorenzahl.[1] Wir ordnen den Zahlen $1, 2, \ldots, m$ die Mengen $A_1, A_2, \ldots, A_m$ zu, bilden also einen Mengenkomplex $(A_1, A_2, \ldots, A_m)$ und ordnen die Menge der zugehörigen Elementenkomplexe $(a_1, a_2, \ldots, a_m)$ lexikographisch, nämlich

$$(a_1, a_2, \ldots, a_m) < (b_1, b_2, \ldots, b_m),$$

wenn entweder

$$a_1 < b_1 \text{ in } A_1,$$

oder

$$a_1 = b_1,\; a_2 < b_2 \text{ in } A_2,$$

oder

$$a_1 = b_1,\; a_2 = b_2,\; a_3 < b_3 \text{ in } A_3 \text{ usw.}$$

Auch der Ausdruck: Ordnung nach den ersten differierenden Elementen oder nach ersten Differenzen ist geeignet, das Wesen dieser Anordnung zu bezeichnen.

Das in dieser Weise geordnete Produkt $\mathfrak{P} A_i$ bezeichnen wir mit

$$A_m A_{m-1} \cdots A_2 A_1$$

und seinen Typus, der nur von den Typen α_i der Mengen A_i abhängt, mit

$$\alpha_m \alpha_{m-1} \cdots \alpha_2 \alpha_1.$$

Das kommutative Gesetz gilt nicht, wohl aber das assoziative, von dem wir uns auf einen Spezialfall beschränken können, aus dem die allgemeine Formel, da es sich nur um endliche Faktorenzahl handelt, abgeleitet werden kann. Es ist nämlich

$$(CB)A \simeq C(BA) \simeq CBA$$

und demnach

$$(\gamma\beta)\alpha = \gamma(\beta\alpha) = \gamma\beta\alpha;$$

denn der Komplex (a, b, c), das geordnete Paar $(p, c) = ((a, b), c)$, dessen erstes Element selbst ein geordnetes Paar $p = (a, b)$ ist, und das

---

[1] Selbst, wenn es sich um eine Folge von Mengen handelt, wäre das lexikographische Verfahren noch anwendbar. Wir verschieben diesen Fall indessen bis zur Entwicklung des allgemeinen Produktbegriffs (Kap. VI, § 3).

geordnete Paar $(p, q) = ((a, b), c)$ entsprechen einander umkehrbar eindeutig, und die lexikographische Anordnung ist in allen drei Fällen dieselbe, wenn a, b, c, p, q resp. die Mengen A, B, C, BA, CB durchlaufen.

Das distributive Gesetz gestattet wiederum nur die Auflösung des letzten Faktors, z. B.

$$CB\sum_j^J A_j = \sum_j^J CBA_j\,.$$

Wir fügen noch die leicht einzusehenden Formeln für Umkehrung einer Summe oder eines Produkts hinzu. Es ist

$$\Big(\sum_j^J A_j\Big)^* = \sum_j^{J^*} A_j^*,$$

z. B.

$$(A_1 + A_2)^* = A_2^* + A_1^*,$$

und

$$(BA)^* = B^*A^*,$$

$$(CBA)^* = C^*B^*A^* \quad \text{usw.}$$

Bei der Summe sind also nicht nur die Summanden, sondern auch das Argument, d. h. die Reihenfolge der Summanden umzukehren; beim Produkt sind nur die Faktoren, nicht ihre Reihenfolge umzukehren.

Auf Grund der Definitionen, die für Summe und Produkt von Mächtigkeiten einerseits, von Ordnungstypen andererseits gegeben worden sind, hat der Typus $\sum \alpha_j$ die Mächtigkeit $\sum \mathfrak{a}_j$, wenn $\mathfrak{a}_j$ die Mächtigkeit des Typus α_j ist, und das Produkt $\beta\alpha$ die Mächtigkeit $\mathfrak{b}\mathfrak{a} = \mathfrak{a}\mathfrak{b}$, wenn $\mathfrak{a}, \mathfrak{b}$ die Mächtigkeiten der Typen α, β sind.

Eine Definition des geordneten Produkts im Falle eines allgemeinen Arguments können wir erst später geben, ebenso auch die Definition einer Potenz mit beliebiger geordneter Basis und geordnetem Exponenten. Indessen hat es natürlich kein Bedenken, schon jetzt $\alpha\alpha = \alpha^2$, $\alpha\alpha\alpha = \alpha^3$ usw. zu schreiben.

Beispiele. Es ist

$$2\omega = 2 + 2 + 2 + \ldots = \omega,$$

$$\omega 2 = \omega + \omega.$$

Ist nämlich $A = \{1, 2, 3, \ldots\}$ die natürliche Zahlenreihe vom Typus ω, $B = \{1, 2\}$ eine Menge von zwei Elementen, also vom Typus 2, so ist $\beta\alpha = 2\omega$ der Typus der lexikographisch geordneten Menge von Paaren (a, b), hingegen $\alpha\beta = \omega 2$ der Typus der lexikographisch ge-

ordneten Menge von Paaren (b, a). Die erste Menge BA besteht aus den Paaren

$$(1,1)\ (1,2)\ (2,1)\ (2,2)\ (3,1)\ (3,2)\ \ldots$$

in der hier angegebenen Reihenfolge (von links nach rechts), die zweite Menge AB aus den Paaren

$$(1,1)\ (1,2)\ (1,3)\ \ldots\ (2,1)\ (2,2)\ (2,3)\ \ldots.$$

Die erste Menge ist der natürlichen Zahlenreihe ähnlich, die zweite nicht, also ist

$$2\omega \neq \omega 2$$

(Ungültigkeit des kommutativen Gesetzes). Eben daraus folgt auch

$$(1+1)\omega = 2\omega = \omega \neq \omega + \omega = 1\omega + 1\omega,$$

ein Beispiel für die Ungültigkeit des distributiven Gesetzes bezüglich des ersten Faktors.

Ebenso ist für jede natürliche Zahl n

$$n\omega = n + n + n + \ldots = \omega,$$
$$\omega n = \omega + \omega + \ldots + \omega,$$

und die Typen ωn sind allesamt verschieden, da eine Menge vom Typus ωn genau n Elemente ohne unmittelbaren Vorgänger (§ 3) besitzt.

Der Typus

$$\omega^2 = \omega\omega = \omega + \omega + \omega + \ldots$$

ist der der lexikographisch geordneten Paarmenge

$$(1,1)\ (1,2)\ (1,3)\ \ldots\ (2,1)\ (2,2)\ (2,3)\ \ldots\ (3,1)\ (3,2)\ (3,3)\ \ldots\ \ldots,$$

also einer Folge (vom Typus ω) von Folgen des Typus ω. Eine Menge dieses Typus ist auch die letzte in der Zusammenstellung S. 72.

Weiter ist z. B.

$$\omega + \omega^2 = \omega(1+\omega) = \omega\omega = \omega^2,$$

während

$$\omega^2 + \omega = \omega(\omega + 1)$$

ein neuer Typus ist.

Nach dem assoziativen Summengesetz ist

$$(\omega+\omega)\omega = (\omega+\omega) + (\omega+\omega) + (\omega+\omega) + \ldots$$
$$= \omega + \omega + \omega + \omega + \ldots = \omega^2,$$

oder nach dem assoziativen Produktgesetz

$$(\omega+\omega)\omega = (\omega 2)\omega = \omega(2\omega) = \omega\omega = \omega^2.$$

Hingegen ist

$$\omega(\omega+\omega) = \omega\omega 2 = \omega^2 2 = \omega^2 + \omega^2$$

ein neuer Typus.

Wir haben hier mehrfach die Verschiedenheit von Typen behauptet, ohne sie ausführlich zu begründen, was gelegentlich bei der Theorie

der Ordnungszahlen (Kap. V) nachgeholt werden wird. Die bisherigen Fälle subsumieren sich größtenteils unter den allgemeineren, daß zwei **Polynome** der Form

$$\omega^n a_0 + \omega^{n-1} a_1 + \dots + \omega^2 a_{n-2} + \omega a_{n-1} + a_n,$$

worin der „Grad“ n eine natürliche Zahl, die „Koeffizienten“ $a_0 \dots a_n$ ganze Zahlen $\geqq 0$ und $a_0 > 0$ ist, nur dann gleich sind, wenn der Grad und die Koeffizienten beiderseits übereinstimmen.

Kehrt man die Menge $\{1, 2, 3, \dots, 0\}$ um, so entsteht $\{0, \dots, 3, 2, 1\}$, also ist

$$(\omega + 1)^* = 1 + \omega^* = 1^* + \omega^*,$$

nicht etwa $= \omega^* + 1$, welches $= (1 + \omega)^* = \omega^*$ ist.

Ebenso ist

$$(\omega + \omega)^* = \omega^* + \omega^*,$$

also

$$(\omega 2)^* = \omega^* 2 = \omega^* 2^*,$$

nicht etwa $= 2^* \omega^* = (2\omega)^* = \omega^*$.

Die Umkehrung von

$$\omega\omega = \omega + \omega + \omega + \dots$$

ist

$$\omega^* \omega^* = \dots + \omega^* + \omega^* + \omega^*.$$

§ 3. Die Strecken einer geordneten Menge.

Jedes Element a der geordneten Menge A bestimmt die Menge A^a der ihm vorhergehenden und die Menge A_a der ihm folgenden Elemente, wobei im Sinne geordneter Addition

$$A = A^a + \{a\} + A_a$$

ist; wir nennen A^a die zu a gehörige **Anfangsstrecke** und A_a die **Endstrecke**. Wenn $A^a = 0$ ist, so ist a das erste Element (Anfangselement) von A und umgekehrt: natürlich braucht A kein erstes Element zu haben. Entsprechendes gilt von A_a und dem etwaigen letzten Element. Das erste oder letzte Element bezeichnen wir als ein **Randelement**; A kann zwei Randelemente oder eins oder keins haben. Eine nichtverschwindende Menge ohne Randelemente heiße **offen**; eine offene Menge ist jedenfalls unendlich.

Zwei Elemente $a < b$ von A bestimmen die Menge

$$A_a^b = \mathfrak{D}(A_a, A^b)$$

der Elemente, die gleichzeitig $> a$ und $< b$ oder **zwischen** a und b liegen; wir nennen sie die zu a, b gehörige **Mittelstrecke**. Es ist

$$A = A^a + \{a\} + A_a^b + \{b\} + A_b.$$

Ist $A_a^b = 0$, so daß also kein Element zwischen a und b liegt, so heißen diese Elemente **benachbart** (konsekutiv), a der **unmittel-**

bare Vorgänger von b, b der unmittelbare Nachfolger von a. Mengen ohne benachbarte Elemente (die Nullmenge und Mengen aus einem Element ausgeschlossen) heißen dicht; eine dichte Menge ist jedenfalls unendlich. Zwischen zwei Elementen einer dichten Menge liegen unendlich viele weitere Elemente.

Die Mengen der reellen oder rationalen oder irrationalen Zahlen in natürlicher Ordnung sind dicht.

Anfangs-, End- und Mittelstrecken heißen zusammen Strecken. Wir rechnen ihnen also die bestimmenden Elemente nicht zu; soll dies geschehen, so werden wir gelegentlich von Intervallen sprechen.

Wir geben die Anfangs- und Endstrecken einer Summe und eines Produkts an. Ist

$$A = \sum_i^J A_i,$$

a_i ein Element von A_i, und gelten die Zerlegungen

$$A_i = B_i + \{a_i\} + C_i,$$
$$J = K + \{i\} + L,$$

(also K die zu i gehörige Anfangsstrecke von J, B_i die zu a_i gehörige Anfangsstrecke von A_i, ebenso L und C_i die bezüglichen Endstrecken), so ist nach dem assoziativen Gesetz

$$A = \sum_k^K A_k + B_i + \{a_i\} + C_i + \sum_l^L A_l,$$

also

$$\sum_k^K A_k + B_i, \quad C_i + \sum_l^L A_l$$

die durch a_i bestimmte Anfangs- resp. Endstrecke von A.

Für ein Produkt

$$A = A_m A_{m-1} \cdots A_2 A_1,$$

d. h. die lexikographisch geordnete Menge der Elementkomplexe

$$a = (a_1, a_2, \ldots, a_{m-1}, a_m)$$

bestimmt sich die durch a bewirkte Zerlegung $A = B + \{a\} + C$ aus den Zerlegungen

$$A_i = B_i + \{a_i\} + C_i \qquad (i = 1, 2, \ldots, m)$$

auf folgende Art, die wir der einfachen Schreibweise wegen an einem Produkt von drei Faktoren illustrieren. Zunächst ist nach dem distributiven Gesetz

$$A_3 A_2 A_1 = A_3 A_2 B_1 + A_3 A_2 \{a_1\} + A_3 A_2 C_1.$$

In der mittleren Menge kann man, da der letzte Faktor aus einem Element besteht, die zugehörigen Komplexe also ein festes An-

fangselement a_1 haben, ersichtlich auf den vorletzten Faktor das distributive Gesetz anwenden und erhält

$$A_3 A_2 \{a_1\} = A_3 B_2 \{a_1\} + A_3 \{a_2\}\{a_1\} + A_3 C_2 \{a_1\},$$

ebenso

$$A_3 \{a_2\}\{a_1\} = B_3 \{a_2\}\{a_1\} + \{a_3\}\{a_2\}\{a_1\} + C_3 \{a_2\}\{a_1\}.$$

Die mittlere Menge der letzten Formel besteht nunmehr aus dem einen Komplex $a = (a_1, a_2, a_3)$; dieser bestimmt also in A die Anfangsstrecke

$$B = A_3 A_2 B_1 + A_3 B_2 \{a_1\} + B_3 \{a_2\}\{a_1\}$$

und die Endstrecke

$$C = C_3 \{a_2\}\{a_1\} + A_3 C_2 \{a_1\} + A_3 A_2 C_1 .$$

Die entsprechende Typenzerlegung ist:

$$\alpha = \alpha_3 \alpha_2 \alpha_1 ,$$
$$\alpha_i = \beta_i + 1 + \gamma_i, \quad \alpha = \beta + 1 + \gamma:$$
$$\beta = \alpha_3 \alpha_2 \beta_1 + \alpha_3 \beta_2 + \beta_3 ,$$
$$\gamma = \gamma_3 + \alpha_3 \gamma_2 + \alpha_3 \alpha_2 \gamma_1 .$$

Es ist leicht, diese Formeln auch unmittelbar aus der lexikographischen Ordnung zu gewinnen.

§ 4. Die Stücke einer geordneten Menge.

Ist a ein Element, M eine nichtverschwindende Teilmenge von A, so drücken wir durch

$$a < M \quad \text{oder} \quad M > a$$

aus, daß a jedem Element m von M vorangeht; analog ist $a > M$ oder $M < a$ zu erklären. Für zwei nichtverschwindende Teilmengen M, N bedeute

$$M < N \quad \text{oder} \quad N > M,$$

daß jedes Element m jedem Element n vorangeht.[1]

Eine nichtverschwindende Teilmenge M von A bestimmt drei weitere Teilmengen von A:

[1] Damit dies nicht zu Kollisionen mit den bereits definierten Elementbeziehungen $a \lesseqgtr b$ führe, wollen wir voraussetzen, daß kein Element von A zugleich Teilmenge von A sei, oder wenigstens, daß von zwei verschiedenen Elementen von A nie zugleich das eine eine Menge und das andere ein Element dieser Menge sei. Sollte dann m zugleich Element und nichtverschwindende Teilmenge M von A sein, so ist $m = M = \{m\}$, und die Relationen $a < m$ und $a < M$ widersprechen einander nicht. Dem Leser wird die gemachte Voraussetzung so selbstverständlich erscheinen, daß er schon ihre Erwähnung für überflüssig hält (vgl. S. 43 Anm.); der Verzicht auf sie würde im folgenden eine umständlichere und weniger suggestive Bezeichnungsweise bedingen.

$$A^M = \text{Menge der Elemente } x < M,$$
$$A_M = \text{Menge der Elemente } z > M,$$
$$|M| = \text{Menge der übrigen Elemente } y.$$

Zu den y gehören die Elemente von M selbst, es ist also $|M|$ von Null verschieden, wogegen eine der ersten beiden Mengen oder beide auch verschwinden können. Soweit diese Elemente vorhanden sind, ist stets $x < y < z$, und es gilt also die Summenformel

$$A = A^M + |M| + A_M,$$

analog der Zerlegung von A durch ein Element.

Ist z. B. A die Menge der reellen Zahlen in natürlicher Ordnung, M die Menge der rationalen Zahlen > 0 und < 1, so ist

A^M die Menge der reellen Zahlen $\leqq 0$,

A_M die Menge der reellen Zahlen $\geqq 1$,

$|M|$ die Menge der reellen Zahlen > 0 und < 1.

Wenn $A^M = 0$, wenn es also kein Element $x < M$ gibt, so könnte man M (analog zu einem Anfangselement) eine Anfangsmenge von A nennen; wir wollen lieber sagen: A ist mit M koinitial. Wenn $A_M = 0$, wenn es also kein Element $z > M$ gibt, so sagen wir: A ist mit M konfinal.[1] Z. B. ist die Menge der reellen Zahlen mit der Menge der natürlichen Zahlen konfinal, weil keine reelle Zahl größer ist als sämtliche natürliche Zahlen, und mit der Menge der ganzen Zahlen sowohl koinitial als konfinal. Alles, was von der einen der beiden Relationen koinitial und konfinal gesagt ist, ist mit den entsprechenden Vertauschungen (z. B. der Zeichen $\lesseqgtr$) auch auf die andere zu übertragen.

Wir können dem Leser den Beweis der Transitivität dieser Relationen überlassen:

I. Ist A mit B, B mit C koinitial, so ist auch A mit C koinitial.

II. Ist $A \geqq B \geqq C$ und A mit C koinitial, so ist auch A mit B und B mit C koinitial.

Wenn M kein letztes, aber A_M ein erstes Element z hat (z also das erste Element $> M$ ist), so schreibt man

$$z = \lim\sup M$$

---

[1] Wir haben also, was der Leser beachten möge, hiermit einseitige Relationen einer Menge zu einer nicht verschwindenden Teilmenge definiert: der Ausdruck „A ist mit M koinitial oder konfinal" setzt stillschweigend $A \geqq M > 0$ voraus. Natürlich könnte man auch eine symmetrische Definition geben: die beiden nichtverschwindenden Teilmengen M, N heißen koinitial, wenn $A^M = A^N$. Obwohl dies den Sprachgebrauch weniger beengen würde, haben wir es vorteilhafter gefunden, bei der Definition des Textes zu bleiben.

und nennt z einen oberen Limes; z ist zugleich oberer Limes der Anfangsstrecke A^z und aller Mengen, mit denen diese konfinal ist. Entsprechend ist ein unterer Limes

$$x = \lim\inf M$$

dadurch definiert, daß A^M ein letztes Element x und M kein erstes Element hat. Die Elemente von A zerfallen hiernach in vier Arten: sie sind entweder nur obere Limites oder nur untere oder beides oder keins von beiden. Z. B. ist in der Menge der reellen oder rationalen Zahlen jedes Element beiderseitiger Limes; in der Menge

$$\left\{\frac{0}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, 1\right\}$$

ist nur 1 oberer Limes.

Zwei nichtverschwindende Teilmengen M, N von A, für die $M < N$, bestimmen die Menge

$$A_M^N = \mathfrak{D}(A_M, A^N)$$

aller Elemente, die zugleich $> M$ und $< N$, also zwischen M und N liegen. Offenbar gilt die Summenformel

$$A = A^M + |M| + A_M^N + |N| + A_N$$

analog der Zerlegung von A durch zwei Elemente; zugleich ist

$$A^N = A^M + |M| + A_M^N,$$
$$A_M = A_M^N + |N| + A_N.$$

Wenn $A_M^N = 0$, also zwischen M und N kein Element liegt, so heißen diese beide Teilmengen benachbart; z. B. sind in der Menge der reellen Zahlen die Menge der rationalen Zahlen < 0 und die Menge der rationalen Zahlen $\geqq 0$ benachbart. Für benachbarte Teilmengen $M < N$ ist

$$A = A^N + A_M.$$

Unter einem Stück einer geordneten Menge A verstehen wir eine Teilmenge M, die zugleich mit zwei Elementen jedes zwischen ihnen liegende Element enthält (wenn also $a < b$ zu M gehören, soll $A_a^b \leqq M$ sein). Wir rechnen zu den Stücken auch die Nullmenge und die aus einem einzigen Element bestehenden Teilmengen.

Ist M ein nichtverschwindendes Stück, so muß jedes Element von $A - M$ entweder $< M$ oder $> M$ sein, wie unmittelbar aus der Definition folgt. Demnach ist

$$A = A^M + M + A_M,$$

die früher definierte Menge $|M|$ ist hier mit M identisch. Zu jedem Stück M ist also eine Summendarstellung $A = P + M + Q$ möglich (auch für $M = 0$, wo diese Darstellung allerdings nicht eindeutig

bestimmt ist); umgekehrt ist evident, daß bei dieser wie bei jeder Summendarstellung $A = \Sigma A_i$ jeder Summand ein Stück ist.

Der Durchschnitt beliebig vieler Stücke ist wieder ein Stück. Die Summe zweier fremder Stücke, deren keins Null ist, ist nur dann ein Stück, wenn die beiden Summanden benachbart sind; die Summe zweier Stücke mit gemeinsamen Elementen ist immer ein Stück. Die Summe beliebig vieler Stücke, die paarweise mindestens ein Element gemein haben, ist jedenfalls ein Stück.

Anfangsstück von A nennen wir eine Teilmenge M, die zu jedem Element gleichzeitig jedes frühere Element enthält (wenn a zu M gehört, soll $A^a \subset M$ sein); entsprechend ist ein Endstück zu erklären. Die Nullmenge betrachten wir sowohl als Anfangs- wie als Endstück. Für ein nichtverschwindendes Anfangsstück resp. Endstück gilt

$$A = M + A_M \quad \text{resp.} \quad A = A^M + M,$$

d. h. eine Summendarstellung $A = M + Q$ resp. $A = P + M$ (auch für $M = 0$ gilt das gleiche); umgekehrt ist der erste resp. letzte Summand einer Summe A ein Anfangsstück resp. Endstück von A.

Anfangs- und Endstücke sind spezielle Fälle von Stücken; Strecken sind spezielle Fälle von Stücken, insbesondere Anfangsstrecken von Anfangsstücken, Endstrecken von Endstücken. Summe und Durchschnitt beliebig vieler Anfangsstücke (Endstücke) ist wieder ein solches. Von zwei verschiedenen Anfangsstücken (Endstücken) ist das eine eine Teilmenge des anderen.

Die oben definierten Mengen A^M, A_M, $|M|$, A_M^N sind sämtlich Stücke von A. Die Menge $|M|$ ist das Stück, das mit M zugleich koinitial und konfinal ist; $A^M + |M|$ ist das Anfangsstück, das mit M konfinal ist, $|M| + A_M$ das mit M koinitiale Endstück. Bei benachbarten Mengen $M < N$ ist A^N mit M konfinal, A_M mit N koinitial.

Wir geben wieder die Anfangs- und Endstücke einer Summe und eines Produkts, d. h. die Zerlegungen

$$A = B + C$$

an (B und C seien von Null verschieden). Der Kürze halber bemerken wir voraus: für eine nichtverschwindende Teilmenge A' von A ist

$$A' = \mathfrak{D}(A', B) + \mathfrak{D}(A', C) = B' + C';$$

sind B', C' beide von Null verschieden, so sagen wir, daß A' zerlegt wird, andernfalls, daß A' unzerlegt bleibt. Ist nun

$$A = \sum_i^J A_i,$$

wobei wir zur Vermeidung nutzloser Weitläufigkeiten alle Sum-

manden > 0 annehmen, so bleibt entweder jeder Summand unzerlegt und dann ist

$$J = K + L,$$

$$B = \sum_k^K A_k, \quad C = \sum_l^L A_l;$$

oder ein Summand A_i (und nur einer) wird zerlegt, und dann ist

$$J = K + \{i\} + L, \quad A_i = B_i + C_i,$$

$$B = \sum_k^K A_k + B_i, \quad C = C_i + \sum_l^L A_l.$$

Für ein Produkt, bei dem wir uns wieder auf drei Faktoren

$$A = A_3 A_2 A_1$$

beschränken wollen, ist zu unterscheiden: entweder bleibt die Menge der Komplexe (a_1, a_2, a_3) mit festem a_1 stets unzerlegt, oder eine dieser Mengen wird zerlegt, aber die Menge der Komplexe mit festem a_1 und a_2 bleibt stets unzerlegt, oder auch eine von diesen Mengen wird zerlegt. Je nachdem haben wir:

entweder

$$A_1 = B_1 + C_1,$$

$$B = A_3 A_2 B_1, \quad C = A_3 A_2 C_1;$$

oder

$$A_1 = B_1 + \{a_1\} + C_1, \quad A_2 = B_2 + C_2,$$

$$B = A_3 A_2 B_1 + A_3 B_2 \{a_1\}, \quad C = A_3 C_2 \{a_1\} + A_3 A_2 C_1;$$

oder

$$A_1 = B_1 + \{a_1\} + C_1, \quad A_2 = B_2 + \{a_2\} + C_2, \quad A_3 = B_3 + C_3,$$

$$B = A_3 A_2 B_1 + A_3 B_2 \{a_1\} + B_3 \{a_2\}\{a_1\},$$

$$C = C_3 \{a_2\}\{a_1\} + A_3 C_2 \{a_1\} + A_3 A_2 C_1.$$

Wir schließen hier noch die Definition der relativen Dichtigkeit an. Die Beziehungen koinitial und konfinal können wir auch so definieren: die (nichtverschwindende) Menge A ist mit M koinitial, wenn es zu jedem Element a von A ein Element m von M derart gibt, daß $m \leqq a$, oder wenn jedes Anfangsintervall $A^a + \{a\}$ mindestens ein Element von M enthält. A ist mit M konfinal, wenn es zu jedem a ein $m \geqq a$ gibt, oder wenn jedes Endintervall $\{a\} + A_a$ mindestens ein Element von M enthält. Nach dieser Analogie definieren wir (allerdings mit verändertem sprachlichem Ausdruck) für eine aus mindestens zwei Elementen bestehende Menge A: die Menge $M \leqq A$ heißt in A dicht, wenn es zu jedem Elementpaar $a < b$ von A ein Elementpaar $m < n$ von M derart gibt, daß $a \leqq m < n \leqq b$, oder wenn jedes Mittelintervall $\{a\} + A_a^b + \{b\}$ mindestens zwei Elemente von M enthält. Hier gelten wiederum die Transitivitätssätze:

III. Ist C in B, B in A dicht, so ist auch C in A dicht.

IV. Ist $A \geqq B \geqq C$ und C in A dicht, so ist auch C in B und B in A dicht.

Für eine (absolut) dichte Menge A können wir einfacher sagen: M ist in A dicht, wenn zwischen zwei Elementen von A stets ein Element von M liegt. Z. B. ist die Menge der rationalen Zahlen in der Menge der reellen Zahlen dicht.

§ 5. Stetigkeit.

Eine Zerlegung in zwei nichtverschwindende Stücke

$$A = P + Q$$

möge ein Sprung heißen, wenn P ein letztes Element und Q ein erstes hat; ein Schnitt, wenn P ein letztes und Q kein erstes, oder umgekehrt P kein letztes und Q ein erstes Element hat; eine Lücke, wenn weder P ein letztes noch Q ein erstes Element hat.

Von zwei benachbarten Teilmengen $M < N$ sagen wir, daß sie die Zerlegung

$$A = A^N + A_M = P + Q$$

bestimmen (P das mit M konfinale Anfangsstück, Q das mit N koinitiale Endstück). Sie bestimmen einen Sprung, wenn M ein letztes Element m und N ein erstes n hat, wobei m, n benachbarte Elemente sind; einen Schnitt, wenn M ein letztes Element $m = \liminf N$ und N kein erstes, oder M kein letztes und N ein erstes Element $n = \limsup M$ hat; eine Lücke, wenn weder M ein letztes noch N ein erstes Element hat.

Eine dichte Menge war eine Menge ohne benachbarte Elemente, d. h. ohne Sprünge. Ihre Zerlegungen sind also Schnitte oder Lücken. Die Menge der rationalen Zahlen z. B. ist dicht; ihre Schnitte sind, wenn r eine rationale Zahl ist,

$$P = A^r + \{r\}, \quad Q = A_r$$

und

$$P = A^r, \quad Q = \{r\} + A_r.$$

Sie hat aber auch Lücken; ist nämlich i eine irrationale Zahl und P die Menge der rationalen Zahlen $< i$, Q die der rationalen Zahlen $> i$, so liefert $P + Q$ eine Lücke. Ähnliches gilt von der Menge der irrationalen Zahlen.

Eine Menge (die Nullmenge und die aus einem Element bestehenden Mengen ausgeschlossen), die weder Sprünge noch Lücken hat, bei der also jede Zerlegung ein Schnitt ist, heißt (im Dedekindschen Sinne) stetig; eine stetige Menge ist also dicht und lückenlos. Die Menge aller reellen Zahlen ist stetig[1]; durch Ein-

[1] Daher die Namen Zahlenkontinuum, Mächtigkeit des Kontinuums.

führung der irrationalen Zahlen in der Dedekindschen Art sind ja eben die Lücken der Menge der rationalen Zahlen ausgefüllt. Dieses Verfahren läßt sich in folgender Weise auf eine beliebige dichte Menge übertragen.

Von zwei verschiedenen Anfangsstücken einer Menge A ist, wie wir schon bemerkten, das eine eine (echte) Teilmenge des anderen; hierdurch erscheinen diese Anfangsstücke selbst geordnet, indem wir von zweien dasjenige als das „frühere" definieren, das Teilmenge des anderen ist. Die so geordnete Menge der Anfangsstücke hat aber stets benachbarte Elemente, nämlich A^a und $A^a + \{a\}$, d. h. Anfangsstrecke und Anfangsintervall von A, die zu einem Element a gehören.

Nehmen wir jetzt A als offene, dichte Menge an und beschränken uns auf diejenigen Anfangsstücke P, die von 0 und A verschieden sind und kein letztes Element haben; diese P bilden eine geordnete Menge $\mathfrak{P}$, von der wir zeigen wollen, daß sie offen, dicht und stetig ist, und daß sie eine in $\mathfrak{P}$ dichte, mit A ähnliche Teilmenge $\mathfrak{A}$ enthält. In der Tat: jedem P entspricht ein von 0 und A verschiedenes Endstück Q ($A = P + Q$); dieses hat entweder ein erstes Element a, in welchem Falle $P = A^a$ eine Anfangsstrecke und $P + Q$ ein Schnitt ist, oder es hat kein erstes Element und $P + Q$ ist eine Lücke. Die schnittbestimmenden P, die Anfangsstrecken, entsprechen vermöge $P = A^a$ umkehrbar eindeutig den Elementen a von A und haben dieselbe Ordnung, bilden also eine mit A ähnliche Menge $\mathfrak{A} \leqq \mathfrak{P}$; die lückenbestimmenden P bilden das Komplement $\mathfrak{B} = \mathfrak{P} - \mathfrak{A}$. Ist nun P eins unserer Anfangsstücke, p ein Element von P, so ist $A^p < P$, woraus folgt, daß $\mathfrak{P}$ kein erstes Element hat (und mit $\mathfrak{A}$ koinitial ist); ist ferner q ein Element von $Q = A - P$ und zwar nicht etwa das erste (Q hat immer unendlich viele Elemente, da A offen ist), so ist $P < A^q$, also hat $\mathfrak{P}$ kein letztes Element (und ist mit $\mathfrak{A}$ konfinal). $\mathfrak{P}$ ist also eine offene Menge. Ist $P_1 < P_2$ und p ein Element von $P_2 - P_1$, aber nicht das erste, so ist $P_1 < A^p < P_2$; dies besagt, daß $\mathfrak{P}$ dicht und $\mathfrak{A}$ in $\mathfrak{P}$ dicht ist. Endlich sei

$$\mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2$$

eine Zerlegung von $\mathfrak{P}$ in zwei nichtverschwindende Stücke, für deren Elemente also durchweg $P_1 < P_2$ gilt. Die Summe P aller Mengen P_1 ist wieder ein Anfangsstück von A, und zwar wie diese von 0 und A verschieden und ohne letztes Element; P ist also Element von $\mathfrak{P}$. Dabei ist, wie unmittelbar ersichtlich, $P_1 \leqq P \leqq P_2$; P ist also, wenn es ein P_1 ist, das letzte, und wenn es ein P_2 ist, das erste. Die Zerlegung $\mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2$ ist demnach jedenfalls keine Lücke; die Menge $\mathfrak{P}$ ist dicht und lückenlos, also stetig.

Wir können diesem Verfahren, freilich mit Preisgabe der eindeutigen Bestimmtheit, eine etwas anschaulichere Fassung geben: wir verschaffen uns eine zu A fremde Menge B, die mit der Menge $\mathfrak{B}$ der lückenbestimmenden Anfangsstücke P äquivalent ist, und übertragen die Ordnung von $\mathfrak{P} = \mathfrak{A} + \mathfrak{B}$[1] auf die äquivalente Menge $A + B$, wobei die eineindeutige Beziehung zwischen A und $\mathfrak{A}$ durch die Zusammengehörigkeit von a und A^a gegeben ist. Die Menge $A + B$ entsteht dann aus A durch Ausfüllung der Lücken, d. h. indem man in jede Lücke $P + Q$ ein Element b einschiebt $(P < b < Q)$, wodurch die b gegenüber den a und auch untereinander geordnet werden.

Ist A die Menge der rationalen Zahlen, so ist B die der irrationalen, $A + B$ die der reellen Zahlen. Was die Elemente b eigentlich „sind", bleibt dabei freilich unentschieden, ist aber auch gleichgültig, da es nur auf ihre Ordnung zu den a und untereinander (und auf ihre Rechengesetze) ankommt; übrigens kann man auch die Anfangsstücke P selber „reelle Zahlen" nennen, die schnittbestimmenden rationale reelle Zahlen, die lückenbestimmenden irrationale reelle Zahlen, und dann, auf die Gefahr einer Verwechselung zwischen a und A^a hin, die Weglassung des Beiworts „reell" verabreden.

§ 6. Dichte, stetige, zerstreute Mengen.

Wir wollen zunächst die Bedingungen aufsuchen, unter denen eine Summe

$$A = \sum_i^J A_i$$

dicht ist. Um wieder triviale Weitläufigkeiten zu vermeiden, nehmen wir die Summanden A_i sämtlich als von Null verschieden an.

Zunächst darf kein Summand Nachbarelemente enthalten; jedes A_i muß also entweder dicht sein oder aus einem einzigen Element bestehen.

Enthält ferner das Argument zwei benachbarte Elemente $i < k$, so darf nicht gleichzeitig A_i ein letztes und A_k ein erstes Element haben.

Diese Bedingungen sind auch hinreichend. Die erste sichert die Existenz eines Elementes zwischen a und b, wenn a, b derselben Menge A_i angehören; die zweite, wenn a und b verschiedenen Mengen A_i, A_k angehören. Beide Bedingungen sind u. a. erfüllt, wenn alle A_i und J dicht sind, oder, bei beliebigem J, wenn alle A_i dicht

[1] Die Summen sind hier nur im Sinne der Addition ungeordneter Mengen zu verstehen; die Elemente von $\mathfrak{A}$ und $\mathfrak{B}$ liegen ja durcheinander.

sind und kein letztes Element haben (gleichviel ob sie ein erstes haben oder nicht), oder wenn sie dicht sind und kein erstes Element haben. Also: Eine Summe dichter Mengen über ein dichtes Argument ist selbst eine dichte Menge; eine Summe dichter Mengen, die kein letztes (erstes) Element haben, ist eine dichte Menge.

Läßt man die Summanden ähnlich werden, so folgt, daß ein (von Null verschiedenes) Produkt BA dicht ist, wenn beide Faktoren dicht sind, oder wenn der erste Faktor B dicht ist und kein letztes (erstes) Element hat. Die Übertragung auf ein Produkt von beliebig vielen Faktoren liegt auf der Hand.

Beispiele. Man bezeichnet den Typus der (natürlich geordneten) Menge der rationalen Zahlen mit η, der reellen Zahlen mit λ. Beide sind offen und dicht, λ auch stetig. Einfache lineare und projektive Transformationen zeigen, daß diese Typen umkehrbar sind ($\eta^* = \eta$, $\lambda^* = \lambda$); daß λ auch der Typus einer Strecke reeller Zahlen ist (der Menge der Zahlen $x < a$, $x > a$, $a < x < b$), η der Typus einer Strecke rationaler Zahlen ($x < a$ usw. für rationale x, a, b). $1 + \lambda$ und $\lambda + 1$ sind die Typen einseitig offener Intervalle,

$$\vartheta = 1 + \lambda + 1$$

der Typus eines Intervalls mit Endpunkten ($a \leqq x \leqq b$ für $a < b$). Durch Aneinanderreihen von Strecken resp. Intervallen erhält man Formeln wie diese

$$\begin{aligned} \lambda &= \lambda + 1 + \lambda, \\ 1 + \lambda &= 1 + \lambda + 1 + \lambda = (1 + \lambda)2 = (1 + \lambda)n \\ &= 1 + \lambda + 1 + \lambda + 1 + \lambda + \ldots = (1 + \lambda)\omega, \\ \lambda + 1 &= (\lambda + 1)2 = (\lambda + 1)n = (\lambda + 1)\omega^*, \\ \lambda &= (\lambda + 1)\omega = (1 + \lambda)\omega^* \end{aligned}$$

und dieselben für η.

Dagegen sind die Typen $\lambda + \lambda = \lambda 2$, λn, $\lambda\omega$ untereinander und von λ verschieden, da sie eine resp. $n - 1$ resp. unendlich viele Lücken haben. Alle diese Typen sind nach unseren allgemeinen Bemerkungen dicht, ebenso $\lambda\eta$, $(1 + \lambda)\eta$, $(\lambda + 1)\eta$, $\vartheta\eta$, λ^2, η^2 usw. Hingegen sind 2λ, $\vartheta 2 = 1 + \lambda + 2 + \lambda + 1$ u. dgl. undicht.

Wir fragen zweitens, wann eine Summe

$$A = \sum_{I}^{J} A_i$$

mit nichtverschwindenden Summanden stetig ist. Nach den Formeln in § 4 gibt es zwei Arten von Zerlegungen $A = B + C$ in zwei nichtverschwindende Stücke. Bei den Zerlegungen der Form

$$J = K + \{i\} + L, \quad A_i = B_i + C_i,$$
$$B = \sum_k^K A_k + B_i, \quad C = C_i + \sum_l^L A_l,$$

wo B_i, C_i von Null verschieden sind, muß, damit $B + C$ ein Schnitt in A sei, $B_i + C_i$ ein Schnitt in A_i sein. Also muß jeder Summand entweder stetig sein oder aus einem einzigen Element bestehen.

Bei den Zerlegungen

$$J = K + L,$$
$$B = \sum_k^K A_k, \quad C = \sum_l^L A_l$$

$(K, L > 0)$ sind folgende Fälle zu unterscheiden:

(α) K hat kein letztes, L kein erstes Element. Dann würde auch B kein letztes, C kein erstes Element haben. Dieser Fall ist also auszuschließen: das Argument muß lückenfrei sein.

(β) K hat ein letztes Element k_0, L kein erstes. Da dann auch C kein erstes Element hat, muß B ein letztes, d. h. A_{k_0} ein letztes Element haben.

(γ) K hat kein letztes Element, L ein erstes l_0. Hier muß A_{l_0} ein erstes Element haben.

(δ) K hat ein letztes Element k_0, L ein erstes l_0. Dann muß A_{k_0} ein letztes und A_{l_0} kein erstes, oder A_{k_0} kein letztes und A_{l_0} ein erstes Element haben.

Diese notwendigen und hinreichenden Bedingungen sind etwas verwickelt. Einfache, hinreichende Bedingungen erhalten wir, wenn wir den Fall (δ) ausschließen: dann ist J stetig (von dem trivialen Fall abgesehen, daß J nur ein Element hat), und da jedes Element i von J die Rolle von k_0 oder l_0 übernehmen kann, muß A_i ein erstes und ein letztes Element haben. Also:

Eine Summe über ein stetiges Argument ist stetig, wenn jeder Summand entweder aus einem Element besteht oder eine stetige Menge mit zwei Randelementen ist.

Lassen wir alle Summanden ähnlich werden, so ergibt sich: ein Produkt BA ist stetig, wenn A und B stetig sind und B ein erstes und ein letztes Element hat.

Z. B. ist der Typus $\vartheta = 1 + \lambda + 1$ stetig, also auch $\vartheta\lambda$, ferner ϑ^2, und da dieser wieder ein erstes und letztes Element hat, auch $\vartheta^2 \cdot \vartheta = \vartheta^3$ und allgemein ϑ^n für jede natürliche Zahl n.

Dagegen ist $\lambda + \lambda = \lambda 2$ unstetig, denn eben diese Zerlegung $\lambda + \lambda$ bestimmt eine Lücke; es ist der Typus der reellen Zahlenmenge nach Weglassung einer Zahl, z. B. der negativen und posi-

tiven Zahlen ohne Null. Auch λ^2 ist unstetig, denn da dies ein offener Typus ist, so bestimmt die Zerlegung

$$\lambda^2 = \lambda(\lambda + 1 + \lambda) = \lambda^2 + (\lambda + \lambda^2)$$

eine Lücke. Auch λ^n ist ein unstetiger Typus.

Wir nennen eine Menge zerstreut, wenn sie keine dichte Teilmenge enthält. Insbesondere sind die endlichen Mengen (zu denen wir auch die Nullmenge rechnen) zerstreut. Jede Teilmenge einer zerstreuten Menge ist selbst zerstreut.

Fragen wir wieder, unter welchen Bedingungen eine Summe

$$A = \sum_i^J A_i$$

mit nichtverschwindenden Summanden zerstreut ist. Die Summanden müssen, als Teilmengen von A, zerstreut sein; desgleichen das Argument, denn A enthält eine mit J ähnliche Teilmenge, bestehend aus je einem Element aller Mengen A_i.

Umgekehrt ist dies auch hinreichend: eine Summe zerstreuter Mengen über ein zerstreutes Argument ist selbst eine zerstreute Menge. Bei dem Beweis dieser Behauptung können wir die Annahme $A_i > 0$ fallen lassen. Soll B eine dichte Teilmenge von

$$A = \sum_i^J A_i$$

sein, wo J und die A_i zerstreut sind, so schreiben wir

$$B_i = \mathfrak{D}(A_i, B), \quad B = \sum_i^J B_i$$

und indem wir nur diejenigen Indices beibehalten, wo $B_i > 0$:

$$B = \sum_k^K B_k.$$

Nach unseren anfänglichen Bedingungen für die Dichtigkeit einer Summe müssen die B_k dicht sein oder nur ein Element enthalten. Da sie als Teilmengen von A_k nicht dicht sein können, so müssen sie aus einem Element bestehen:

$$B = \sum_k^K \{b_k\} \simeq K.$$

B kann also nicht dicht sein, da K nicht dicht ist, als Teilmenge von J. Damit ist die Behauptung bewiesen.

Wir können hiermit einen für den Aufbau geordneter Mengen wichtigen Satz beweisen, nämlich:

Jede geordnete Menge ist entweder zerstreut oder die Summe zerstreuter Mengen über ein dichtes Argument.

A sei eine beliebige geordnete Menge. Wir wollen einmal vorübergehend ein Element a mit b **kongruent** nennen ($a \equiv b$), wenn

entweder $b = a$

oder, für $b > a$, die Mittelstrecke A_a^b zerstreut ist

oder, für $b < a$, die Mittelstrecke A_b^a zerstreut ist.[1]

Zunächst folgt aus $a \equiv b$ auch $b \equiv a$.

Ist ferner $a \equiv b$, $b \equiv c$, so ist auch $a \equiv c$. Um das zu zeigen, können wir alle drei Elemente als verschieden annehmen, ferner $a < c$ (indem wir andernfalls alle drei Kongruenzen umdrehen); für b sind dann die drei Lagen

$$b < a < c, \quad a < b < c, \quad a < c < b$$

möglich. Im ersten und dritten Fall schließen wir, daß A_a^c als Teilmenge von A_b^c resp. A_a^b zerstreut ist. Im zweiten Fall ist die Summeneigenschaft zu benutzen und zu schließen, daß

$$A_a^c = A_a^b + \{b\} + A_b^c$$

als Summe dreier zerstreuter Mengen wieder zerstreut ist.

Wir vereinigen jetzt alle mit einem gegebenen Element a kongruenten Elemente zu einer Menge $A(a)$, die jedenfalls das Element a selbst enthält. Ist b irgend ein von a verschiedenes Element von $A(a)$, so ist $A(b) = A(a)$, denn es ist $a \equiv b$, und alle mit b kongruenten Elemente sind auch mit a kongruent und umgekehrt. Zwei solche Mengen, die ein Element gemein haben, sind also identisch.

Jede Menge $A(a)$ ist selbst eine **zerstreute** Menge. Denn enthielte sie eine dichte Teilmenge B, und sind $b < c$ zwei Elemente von B, so wäre auch B_b^c dicht, während dies doch eine Teilmenge der zerstreuten Menge A_b^c ist.

Endlich ist $A(a)$ ein **Stück** von A. Denn sind $b < c$ zwei ihrer Elemente, d ein Element zwischen beiden, so ist A_b^d als Teilmenge von A_b^c zerstreut, also $d \equiv b \equiv a$; d gehört zu $A(a)$.

Zwei verschiedene Mengen $A(a)$, $A(b)$ haben, wie oben bemerkt, kein Element gemein; ist $a < b$, so gehen alle Elemente von $A(a)$ allen von $A(b)$ voran $\big(A(a) < A(b)\big)$. Nehmen wir aus jeder Menge $A(a)$ je ein Element heraus, so bilden diese eine Teilmenge J von A, und es ist

$$A = \sum_i^J A(i).$$

Entweder besteht nun J aus einem einzigen Element, dann ist A selbst zerstreut. Wenn nicht, so kann J keine Nachbarelemente $i < k$

[1] Insbesondere dürfen diese Mittelstrecken auch Null sein.

enthalten; sonst wäre nämlich $A(i) + A(k)$ zerstreut, also auch A_i^k, folglich $i \equiv k$, während doch i, k inkongruent vorausgesetzt waren. In diesem Fall ist also J dicht, A eine Summe zerstreuter Mengen mit dichtem Argument.

§ 7. Abzählbare Typen.

Die Menge der Typen einer bestimmten Mächtigkeit $\mathfrak{a}$ nennen wir eine Typenklasse und bezeichnen sie mit $T(\mathfrak{a})$. Die Typenklassen endlicher Mächtigkeit bestehen aus nur einem Typus: $T(0)$ aus dem Typus 0, $T(n)$ aus dem Typus n. Von der Klasse $T(\aleph_0)$ der abzählbaren Typen kennen wir schon viele verschiedene Vertreter; wir zeigen jetzt, daß sie unendlich viele Typen enthält und zwar, daß sie von der Mächtigkeit des Kontinuums ist.

Zunächst ist jede Typenklasse $T(\mathfrak{a})$ von der Mächtigkeit $\leqq 2^{\mathfrak{a}\mathfrak{a}}$. Betrachten wir nämlich der Reihe nach die Mengen mit folgenden Elementen:

(α) die verschiedenen Typen der Mächtigkeit $\mathfrak{a}$;

(β) die im Typus verschiedenen Ordnungen einer Menge A von der Mächtigkeit $\mathfrak{a}$;

(γ) alle Ordnungen der Menge A;

(δ) alle Paarmengen (§ 1) P, die eine Ordnung von A hervorrufen;

(ε) alle Paarmengen P aus geordneten Paaren (a, b) mit $a \neq b$;

(ζ) alle Paarmengen P aus geordneten Paaren (a, b).

Hier ist $\alpha \sim \beta \leqq \gamma \sim \delta \leqq \varepsilon \leqq \zeta$, und ζ ist das System der Teilmengen der Menge aller geordneten Paare. Diese Menge aller Paare hat die Mächtigkeit $\mathfrak{a}\mathfrak{a} = \mathfrak{a}^2$, die Menge ζ also die Mächtigkeit $2^{\mathfrak{a}\mathfrak{a}}$, und unsere gesuchte Menge eine Mächtigkeit $\leqq 2^{\mathfrak{a}\mathfrak{a}}$.

Die Klasse $T(\aleph_0)$ hat demnach eine Mächtigkeit

$$\leqq 2^{\aleph_0 \aleph_0} = 2^{\aleph_0} = \aleph.$$

Wir zeigen, daß sie auch eine Mächtigkeit $\geqq \aleph$ hat, also genau $\aleph$.

Zu diesem Zweck betrachten wir etwa, wenn $a = (a_1, a_2, \ldots)$ eine Folge natürlicher Zahlen ist, den Typus

$$\alpha = \sum_i (a_i + \eta) = a_1 + \eta + a_2 + \eta + a_3 + \eta + \cdots$$

der Menge

$$A = \{1, 2, \ldots, a_1\} + R_1 + \{a_1 + 1, \ldots, a_1 + a_2\} + R_2 + \{a_1 + a_2 + 1, \ldots, a_1 + a_2 + a_3\} + R_3 + \cdots,$$

wobei alle R_i der Menge der rationalen Zahlen ähnlich und untereinander sowie mit der Menge J der natürlichen Zahlen fremd sein sollen. A ist also eine bestimmte Anordnung der aus J und

$$R = R_1 + R_2 + R_3 + \cdots$$

gebildeten Summe $J+R$ (aber nicht in der Ordnung $J+R$ oder $R+J$).

Wir behaupten, daß die Zahlenfolgen a und diese Mengen A oder ihre Typen α einander umkehrbar eindeutig entsprechen, d. h. daß zwei verschiedenen Zahlenfolgen auch verschiedene Typen entsprechen, oder daß A mit

$$B = \{1, 2, \ldots, b_1\} + R_1 + \{b_1 + 1, \ldots, b_1 + b_2\} + R_2 + \cdots$$

nur dann ähnlich sein kann, wenn $a_1 = b_1$, $a_2 = b_2$, $a_3 = b_3, \ldots$. Nehmen wir also die beiden geordneten Mengen A und B, die ja aus denselben Elementen bestehen, ähnlich an. Die Elemente von R haben sowohl in A als auch in B kein benachbartes Element, während die Elemente von J in beiden geordneten Mengen mindestens ein benachbartes Element haben. Die Abbildung $b = f(a)$ kann also nur so geschehen, daß die Mengen R, J einzeln auf sich selber abgebildet werden, d. h. $f(r)$ ein r und $f(i)$ ein i ist. Auf Grund der Ähnlichkeit ist die letzte Beziehung nur so möglich, daß

$$f(1) = 1,\; f(2) = 2,\; \ldots,\; f(i) = i,\; \ldots$$

ist.

Andererseits müssen die Elemente von J, die in A keinen unmittelbaren Nachfolger haben, auf ebensolche in B abgebildet werden. Diese Elemente sind, der Reihe nach,

$$\text{in } A\colon\; a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots,$$
$$\text{in } B\colon\; b_1, b_1 + b_2, b_1 + b_2 + b_3, \ldots.$$

Also folgt:

$$b_1 = f(a_1) = a_1,$$
$$b_1 + b_2 = f(a_1 + a_2) = a_1 + a_2,\; b_2 = a_2$$

usw.

Hiernach enthält, da die Typen α abzählbar sind, $T(\aleph_0)$ eine mit der Menge aller Folgen natürlicher Zahlen äquivalente Teilmenge; demnach ist die Mächtigkeit dieser Typenklasse $\geqq \aleph_0^{\aleph_0} = \aleph$ und mithin genau $= \aleph$.

Natürlich könnte man statt der obigen Typen α andere wählen; wir laden den Leser ein, es z. B. mit

$$\sum_i (a_i + \omega^* + \omega) = a_1 + \omega^* + \omega + a_2 + \omega^* + \omega + \cdots$$

zu versuchen.

A b z ä h l b a r e d i c h t e T y p e n. Wir kennen deren vier, nämlich

$$\eta,\; 1 + \eta,\; \eta + 1,\; 1 + \eta + 1,$$

und es ist nun eine sehr bemerkenswerte Tatsache (zu der für höhere Mächtigkeiten kein Analogon oder nur ein sehr beschränktes existiert), daß es weitere abzählbare dichte Typen nicht gibt, daß

insbesondere jeder offene abzählbare dichte Typus $= \eta$ ist. Wir beweisen nämlich die beiden Sätze:

I. Eine offene dichte Menge enthält zu jeder abzählbaren Menge eine ähnliche Teilmenge.

II. Zwei offene dichte abzählbare Mengen sind ähnlich.

B sei eine offene dichte Menge, A eine beliebig geordnete abzählbare Menge, die wir also in Gestalt einer Folge

$$A = \{a_1, a_2, a_3, \ldots\}$$

bringen können, wobei die Rangordnung der a_i in A aber nicht die der Indices zu sein braucht.

Dem a_1 ordnen wir ein beliebiges Element b_1 von B zu. Dem a_2 müssen wir ein Element b_2 zuordnen, das $\lesseqgtr b_1$ ist, je nachdem $a_2 \lesseqgtr a_1$ ist; solche Elemente b_2 existieren sicher, da b_1 weder das erste noch das letzte Element von B ist, und wir denken uns eins von ihnen gewählt. Sodann ist dem a_3 ein b_3 zuzuordnen, das zu b_1, b_2 dieselbe Rangordnung hat wie a_3 zu a_1, a_2; auch solche b_3 existieren sicher, sei es, daß a_3 zwischen a_1 und a_2 liegt oder beiden vorangeht oder beiden nachfolgt, da ja in B kein Element das erste oder letzte und keine zwei Elemente benachbart sind.

Man sieht hier schon den Umstand, auf den es ankommt und der die unbegrenzte Fortsetzung des Verfahrens sichert. Angenommen, wir hätten bereits zu $a_1, a_2, \ldots, a_n$ passende Bilder $b_1, b_2, \ldots, b_n$ bestimmt. Das nächste Element a_{n+1} liegt entweder zwischen zweien der Elemente $a_1, a_2, \ldots, a_n$ oder geht allen voraus oder folgt allen nach: in jedem Fall kann man ein Element b_{n+1} in entsprechender Lage zu den $b_1, b_2, \ldots, b_n$ wählen, da B kein erstes und kein letztes Element und keine benachbarten Elemente hat. Demnach ist es möglich, jedem Element a_i ein passendes b_i zuzuordnen: A ist mit einer Teilmenge von B ähnlich (I).

Um weiter II zu beweisen, nehmen wir zwei offene, dichte, abzählbare Mengen

$$A = \{a_1, a_2, a_3, \ldots\}$$
$$B = \{b_1, b_2, b_3, \ldots\}$$

(die Indices der b haben jetzt nicht dieselbe Bedeutung wie soeben) und wenden das geschilderte Verfahren so an, daß wir abwechselnd einem a_i ein b_k und einem b_k ein a_i zuordnen, dabei aber durch Vorgehen in der Reihenfolge der Indices dafür sorgen, daß von beiden Mengen kein Element übergangen wird. Wir ordnen also dem $a_1 = a'$ irgend ein Element b' zu; sodann sei b'' unter den noch freien b_k, d. h. in der Menge $B - \{b'\}$, das mit niedrigstem Index (z. B. wenn $b' = b_1$ gewählt war, $b'' = b_2$; in jedem anderen

Falle $b'' = b_1$), und ihm ordnen wir ein a'' zu, das zu a' dieselbe Ordnung hat wie b'' zu b'. Dann sei a''' unter den noch freien a_i, d. h. in der Menge $A - \{a', a''\}$, das mit niedrigstem Index und b''' ein passendes Bild dazu, danach wieder b'''' das niedrigste noch unabgebildete b_k und a'''' ein passendes Bild usw. Die Existenz passender Bilder[1] ist durch die Offenheit und Dichtigkeit beider Mengen verbürgt, und die Wahl der niedrigsten Indices schließt das Übergehen eines Elements aus; man sieht ja, daß nach den ersten $2n$ Schritten die Elemente $a_1, \ldots, a_n$ und $b_1, \ldots, b_n$ sicher abgebildet sind. Also ist A mit B ähnlich, womit II bewiesen ist.

Da nun jede dichte Menge nach Abtrennung etwaiger Randelemente eine offene dichte Menge übrig läßt, so gibt es nur die vier oben aufgeführten abzählbaren dichten Typen.

Daraus folgen Formeln wie diese:

$$\eta = \eta + \eta = \eta 2 = \eta n$$
$$= \eta + \eta + \eta + \ldots = \eta\omega.$$

Allgemein ist $\eta\alpha = \eta$, wenn α ein abzählbarer Typus ist. Denn alle diese Typen sind nach § 6 dicht, ferner offen und abzählbar.

Ebenso ist, wie wir schon wissen,

$$\eta = \eta + 1 + \eta,$$
$$1 + \eta = (1 + \eta)2 = (1 + \eta)n = (1 + \eta)\omega;$$

allgemein ist $(1 + \eta)\alpha = 1 + \eta$ oder $= \eta$, je nachdem α ein erstes Element hat oder nicht.

$(1 + \eta + 1)n$ ist nur für $n = 1$ dicht; $(1 + \eta + 1)\alpha$ nur, wenn α dicht ist, und zwar ist dann $(1 + \eta + 1)\alpha = \alpha$.

Wir werden später zu den Sätzen I, II zwar ein Analogon für höhere Mächtigkeiten kennen lernen (Kap. VI, § 8), aber nur unter Hinzufügung weiterer Forderungen außer der Dichtigkeit. Daß es z. B. in der Typenklasse $T(\aleph)$ sicher unendlich viele offene dichte Typen gibt, lehren schon die Beispiele

$$\lambda,\ \lambda 2 = \lambda + \lambda,\ \lambda 3 = \lambda + \lambda + \lambda,\ \ldots,$$

von denen der erste Typus keine, der zweite eine, der dritte zwei Lücken enthält usw.

Abzählbare stetige Typen gibt es überhaupt nicht; denn stetige Typen sind ja dicht, und die vier abzählbaren dichten Typen sind unstetig.

Wir können sogar noch mehr sagen. Eine offene dichte

---

[1] Man kann auch für diese noch festsetzen, daß man unter den in Betracht kommenden a_i, b_k immer das mit niedrigstem Index wählen solle, und dadurch die Abbildung vollkommen fixieren.

Menge A enthält nach I eine Teilmenge B, die mit der Menge der rationalen Zahlen ähnlich ist. Ist $B = P + Q$ eine Zerlegung von B, die eine Lücke definiert, so muß, wenn A stetig sein soll, zwischen P und Q mindestens ein Element von A liegen; denn wenn P, Q in A benachbart wären, würden sie auch dort eine Lücke $A = A^Q + A_P$ bestimmen. A enthält also eine Teilmenge C, die aus B durch Ausfüllung der Lücken entsteht, d. h. mit der Menge der reellen Zahlen ähnlich ist. Andererseits sieht man noch: wenn B in A dicht ist, so kann zwischen P und Q auch nur höchstens ein Element von A liegen. Das gibt den Satz:

III. Jede stetige Menge enthält eine Teilmenge, die mit der Menge der reellen Zahlen ähnlich ist. Jede offene stetige Menge, in der eine abzählbare Menge dicht ist, ist mit der Menge der reellen Zahlen ähnlich.

Die Typen $\vartheta = 1 + \lambda + 1$ und ϑ^2 sind, wie wir sahen (S. 94), beide stetig, aber verschieden. In ϑ ist nämlich eine abzählbare Teilmenge dicht, in ϑ^2 gibt es keine solche. Denn ϑ^2 ist der Typus der lexikographisch geordneten Menge von reellen Zahlenpaaren (a, b); eine in dieser Menge dichte Teilmenge muß für jedes a mindestens ein Paar enthalten (sogar eine solche Menge von Paaren, die in der Menge aller Paare mit diesem bestimmten a dicht ist), ist also von der Mächtigkeit des Kontinuums. Der Beweis, daß alle Potenzen ϑ^n verschiedene Typen sind, ist nicht ganz so einfach.

Fünftes Kapitel.

Wohlgeordnete Mengen. Ordnungszahlen.

§ 1. Wohlordnung.

Bei dem Versuch (Kap. IV, § 1), die Eigenschaften der natürlichen Zahlenreihe auf unendliche Mengen zu übertragen, haben wir zunächst das Moment der Ordnung berücksichtigt. Die Zahlenreihe ist aber eine sehr spezielle geordnete Menge, und ihre Funktion als Instrument zum Zählen knüpft sich gerade an eine solche spezielle Eigenschaft, daß nämlich, wenn man bis n gezählt hat, nunmehr eine nächstfolgende Zahl $n + 1$ an die Reihe kommt. Anders ausgedrückt: wir haben hier eine geordnete Menge A von der Beschaffenheit, daß bei jeder Zerlegung $A = P + Q$ das Endstück Q (falls es Elemente enthält) ein erstes Element hat. Diese Eigenschaft übertragen wir auf beliebige Mengen. Eine geordnete Menge A

heiße wohlgeordnet und ihr Ordnungstypus eine Ordnungszahl, wenn jedes von Null verschiedene Endstück ein erstes Element hat.

Wir können auch sagen: A ist wohlgeordnet, wenn jede von Null verschiedene Teilmenge A' ein erstes Element hat. Diese Bedingung ist ja offenbar hinreichend, aber auch notwendig; denn (Kap. IV, § 4) A' bestimmt das mit ihm koinitiale Endstück, dessen erstes Element auch das erste Element von A' ist. Es ist in der Definition eingeschlossen, daß auch A selbst ein erstes Element haben muß. Jedes Anfangsstück der wohlgeordneten Menge ist, wenn nicht A selbst, eine Anfangsstrecke, denn dann hat $A - A'$ ein erstes Element a und es ist $A' = A^a$.

Jede endliche Menge ist wohlgeordnet; wir betrachten auch die Nullmenge als wohlgeordnet. $0, 1, 2, \ldots$ sind also Ordnungszahlen.

Die natürliche Zahlenreihe ist wohlgeordnet, also auch ω ist eine Ordnungszahl. Das Gleiche gilt von $\omega + n$, $\omega + \omega$ u. dgl., während die Typen ω^*, η, λ keine Ordnungszahlen sind. In einer wohlgeordneten Menge gibt es keine Teilmenge vom Typus ω^* (auch dies könnte als Definition genommen werden).

Jede Teilmenge einer wohlgeordneten Menge ist wohlgeordnet.

Eine Summe wohlgeordneter Mengen über ein wohlgeordnetes Argument ist wohlgeordnet. Denn ist

$$A = \sum_i^J A_i,$$

wo J und die A_i wohlgeordnet sind, so hat A, falls > 0, ein erstes Element. Ist nämlich K die Menge der Indices i, für die $A_i > 0$, so ist

$$A = \sum_k^K A_k;$$

nun hat K ein erstes Element k_0 (als Teilmenge von J) und A_{k_0} ein erstes Element, das denn auch das erste Element von A ist. Jede Teilmenge A' von A ist wieder eine solche Summe

$$A' = \sum_i^J \mathfrak{D}(A_i, A') = \sum_i^J A_i',$$

wo J und die $A_i' (\leqq A_i)$ wohlgeordnet sind, und hat also, falls > 0, ein erstes Element; demnach ist A wohlgeordnet.

Läßt man die Summanden ähnlich werden, so folgt: ein Produkt wohlgeordneter Mengen ist wohlgeordnet.

Eine Summe von Ordnungszahlen mit wohlgeordnetem Argument und ein Produkt von Ordnungszahlen ist also wieder eine Ordnungszahl. Beispiele: $\omega + n$, $\omega + \omega$, ωn, $\omega\omega = \omega^2$, ω^n,

$$\sum_n \omega^n = \omega + \omega^2 + \omega^3 + \ldots \text{ usw.}$$

§ 2. Die Vergleichbarkeit der Ordnungszahlen.

Jedes Element x einer von 0 verschiedenen wohlgeordneten Menge A bestimmt die Anfangsstrecke

$$X = A^x$$

der Elemente $< x$ und das Endintervall

$$A - X = \{x\} + A_x$$

der Elemente $\geqq x$; wir nennen hier (im Anschluß an die Cantorschen Bezeichnungen) X auch den durch x bestimmten Abschnitt, $A - X$ den durch x bestimmten Rest von A. Es ist stets $A \geqq X$. Ist x das erste Element von A, so ist $X = 0$, $A - X = A$.

I. Eine wohlgeordnete Menge ist niemals einem ihrer Abschnitte ähnlich.

Dieser Satz ist ein Spezialfall des folgenden:

II. Ist $a' = f(a)$ eine ähnliche Abbildung der wohlgeordneten Menge A auf eine ihrer Teilmengen A', so kann kein Element auf ein früheres abgebildet werden, d. h. niemals $f(a) < a$ sein.

Gäbe es nämlich solche Elemente, für die $f(a) < a$ wäre, so würde die Menge dieser Elemente, als von 0 verschiedene Teilmenge von A, ein erstes Element a haben. Danach wäre

$$a' = f(a) < a, \quad a' < a,$$

wegen der Ähnlichkeit der Abbildung also

$$f(a') < f(a), \quad f(a') < a'$$

und im Widerspruch zur Annahme gäbe es vor a ein noch früheres Element a' von der genannten Eigenschaft.

Aus dem hiermit bewiesenen Satz II folgt I unmittelbar; denn sollte $a' = f(a)$ eine ähnliche Abbildung von A auf $A' = A^x$ sein, so müßte ja $f(x)$ zu A' gehören, also $f(x) < x$ sein.

Wir definieren jetzt für zwei Ordnungszahlen α, β die Beziehung

$$(1) \qquad \alpha < \beta \quad \text{oder} \quad \beta > \alpha$$

(α ist kleiner als β, β ist größer als α), wenn die Menge A vom Typus α einem Abschnitt der Menge B vom Typus β ähnlich ist. Offenbar kommt es auf die Mengen A, B selber nicht an: sie können durch ähnliche ersetzt werden.

Aus $\alpha < \beta$, $\beta < \gamma$ folgt $\alpha < \gamma$; denn ist A einem Abschnitt von B, B einem Abschnitt von C ähnlich, so ist A einem Abschnitt eines Abschnittes von C, also einem Abschnitt von C ähnlich. Das Zeichen $<$ (und ebenso $>$) befolgt das transitive Gesetz.

Aus I folgt nun sofort, daß niemals $\alpha < \alpha$ ist, d. h. die Relationen $\alpha < \beta$ und $\alpha = \beta$ schließen einander aus. Ebenso schließen einander die Relationen $\alpha > \beta$ und $\alpha = \beta$ aus.

Ferner schließen die Relationen $\alpha < \beta$ und $\alpha > \beta$ einander aus; denn aus $\alpha < \beta$, $\beta < \alpha$ würde $\alpha < \alpha$ folgen (nach dem transitiven Gesetz).

Von den drei Relationen zwischen Ordnungszahlen

$$\alpha \gtreqless \beta$$

kann also höchstens eine eintreten. Wir zeigen jetzt, daß auch eine davon eintreten muß, daß also zwei Ordnungszahlen stets vergleichbar sind.

Wir führen hier eine im folgenden durchweg festgehaltene Bezeichnung ein:

$$W(\alpha) = \text{Menge der Ordnungszahlen} < \alpha.$$

Zunächst zeigen wir, daß jede wohlgeordnete (von Null verschiedene) Menge A vom Typus α mit $W(\alpha)$ ähnlich ist. Jedem Element x von A entspricht ein Abschnitt X und eine Ordnungszahl $\xi < \alpha$, wobei, für $x < x'$, x ein Element von X', also X ein Abschnitt von X' und $\xi < \xi'$ ist. Andererseits entspricht jeder Ordnungszahl ξ, die kleiner als α ist, mindestens ein Abschnitt X von A, und nach dem soeben Festgestellten auch nicht mehr als einer. Durch die Zuordnung zwischen x und dem Typus ξ des zugehörigen Abschnitts wird also die Menge A auf die Menge $W(\alpha)$ aller Ordnungszahlen $< \alpha$ ähnlich abgebildet, d. h. $W(\alpha)$ ist wohlgeordnet und vom Typus α. Zu beachten ist, daß zu $W(\alpha)$ die Zahl 0, als Typus des zum ersten Element von A gehörigen Abschnitts, mitzurechnen ist. Diese Zuordnung berechtigt uns, die Menge A mit Hilfe von Indices in der Gestalt

$$A = \{a_0, a_1, \ldots, a_\xi, \ldots\}$$

zu schreiben, wobei jeder Zahl $\xi (< \alpha)$ eindeutig umkehrbar ein Element a_ξ entspricht und die Ordnung der Elemente in A mit der Größenordnung ihrer Indices identisch ist ($a_\xi \lesseqgtr a_\eta$ für $\xi \lesseqgtr \eta$).

Für eine natürliche Zahl n ist $W(n) = \{0, 1, \ldots, n-1\}$, denn dem p^{ten} Element der Menge A $(p = 1, 2, \ldots, n)$ entspricht als Abschnitt die Menge der $p - 1$ vorhergehenden Elemente, also die Ordnungszahl $p - 1$. $W(n)$ ist vom Typus n. Ebenso ist

$$W(\omega) = \{0, 1, 2, \ldots\},$$

denn ist A vom Typus ω, etwa die natürliche Zahlenreihe, so entspricht ihrem p^{ten} Element wieder die Ordnungszahl $p - 1$. $W(\omega)$ ist vom Typus ω.

$W(0)$ ist $= 0$ zu setzen, d. h. die Nullmenge, während $W(1)$ die aus der einen Zahl 0 bestehende Menge $\{0\}$ ist.

Mit Hilfe der Mengen $W(\alpha)$ ist nun der Vergleichbarkeitsbeweis einfach zu führen. Jeder Abschnitt von $W(\alpha)$ ist durch ein Element ξ von $W(\alpha)$ bestimmt, besteht also aus allen Zahlen $< \alpha$, die zugleich

$<\xi$ sind, also wegen $\xi<\alpha$ aus allen Zahlen $<\xi$, und ist folglich mit $W(\xi)$ identisch. Jedes Anfangsstück von $W(\alpha)$ ist $W(\alpha)$ selbst oder ein Abschnitt, d. h. ist ein $W(\xi)$ mit $\xi\leqq\alpha$. Dies vorausgeschickt, seien α und β zwei Ordnungszahlen. Der Durchschnitt $D=\mathfrak{D}(W(\alpha), W(\beta))$ ist ein Anfangsstück beider Mengen, da zugleich mit einer Zahl ξ auch jede Zahl $<\xi$ zu D gehört. Demnach ist $D=W(\delta)$ mit $\delta\leqq\alpha$, $\delta\leqq\beta$. Das gibt die vier Fälle:

$$\delta=\alpha,\quad \delta=\beta : \alpha=\beta,$$
$$\delta<\alpha,\quad \delta=\beta : \alpha>\beta,$$
$$\delta=\alpha,\quad \delta<\beta : \alpha<\beta,$$
$$\delta<\alpha,\quad \delta<\beta.$$

Aber der vierte Fall ist ausgeschlossen, weil dann δ zugleich zu $W(\alpha)$ und $W(\beta)$, also zu $D=W(\delta)$ gehören würde, während diese Menge nur die Zahlen $<\delta$ umfaßt.

Damit ist der Vergleichbarkeitssatz bewiesen:

III. Zwei Ordnungszahlen α, β sind stets vergleichbar, d. h. es besteht zwischen ihnen eine und nur eine der drei Relationen

$$\alpha<\beta,\quad \alpha=\beta,\quad \alpha>\beta.$$

Ist insbesondere $A\leqq B$, so ist $\alpha\leqq\beta$. Dann wäre $\alpha>\beta$, B also einem Abschnitt B' von A ähnlich, der durch das Element x bestimmt ist, so würde bei der ähnlichen Abbildung $b'=f(b)$ von B auf B' im Widerspruch zu II $f(x)<x$ sein.

Natürlich kann $A<B$ und dennoch $\alpha=\beta$ sein, da A keinem Abschnitt von B ähnlich zu sein braucht. Z. B. haben die unendlichen Teilmengen der natürlichen Zahlenreihe alle den Typus ω.

IV. Jede Menge von Ordnungszahlen ist, wenn diese Zahlen der Größe nach geordnet werden, eine wohlgeordnete Menge.

Eine nichtverschwindende Menge W von Ordnungszahlen hat ein erstes Element; denn ist α eine zu W gehörige Zahl und noch nicht die kleinste, so hat $\mathfrak{D}(W, W(\alpha))$ als von Null verschiedene Teilmenge der wohlgeordneten Menge $W(\alpha)$ eine kleinste Zahl, und diese ist die kleinste Zahl von W. Da von jeder Teilmenge von W das Gleiche gilt, so ist W wohlgeordnet und kann, falls sie vom Typus β ist, in der Gestalt

$$W=\{\alpha_0, \alpha_1, \ldots, \alpha_\eta, \ldots\} \quad (\eta<\beta) \tag{2}$$

geschrieben werden, wobei die Größenordnung der Zahlen α_η zugleich die ihrer Indices ist.

Zu jeder Ordnungszahl gibt es größere. Sind A, C wohlgeordnet

und fremd, C von Null verschieden, so ist A ein Abschnitt von $A + C$, also

$$\alpha + \gamma > \alpha \quad \text{für} \quad \gamma > 0;$$

umgekehrt ist jede Zahl $\beta > \alpha$ in dieser Form $\beta = \alpha + \gamma$ darstellbar, denn ist A einem Abschnitt A' von B ähnlich, so ist $B = A' + C$. Da C mindestens ein Element haben muß, so ist $\alpha + 1$ die nächstgrößere Zahl über α.

Auch zu jeder Menge W von Ordnungszahlen gibt es größere Ordnungszahlen.[1] Für die Menge (2) ist z. B.

$$\alpha = \sum_\eta (\alpha_\eta + 1)$$

eine Zahl $> W$; denn es ist, für jedes η, $\alpha \geqq \alpha_\eta + 1 > \alpha_\eta$. Unter den Zahlen $> W$ gibt es eine kleinste, denn wenn α noch nicht die kleinste ist, so hat die Menge der Zahlen von $W(\alpha)$, die $> W$ sind, eine kleinste Zahl, und diese ist die gesuchte. Der Leser wird übrigens leicht feststellen können, daß die kleinste Zahl $\sigma > W$ der Typus der Menge ist

$$W(\sigma) = \mathfrak{S}_\eta W(\alpha_\eta + 1).$$

Eine Ordnungszahl $\lambda > 0$ ohne unmittelbaren Vorgänger nennt man eine Limeszahl; sie ist der Typus einer nichtverschwindenden Menge $W(\lambda)$ ohne letztes Element. Die kleinste Limeszahl ist ω. Hat die nichtverschwindende Menge W von Ordnungszahlen kein letztes Element, so ist die nächstgrößere Zahl eine Limeszahl, die man mit[2]

$$\lambda = \lim W$$

oder bei der Schreibart (2) auch mit

$$\lambda = \lim_\eta \alpha_\eta = \overset{W(\beta)}{\lim_\eta} \alpha_\eta$$

bezeichnet (wobei auch β eine Limeszahl ist). Jede Limeszahl λ ist Limes von $W(\lambda)$ und denjenigen Mengen, mit denen $W(\lambda)$ konfinal ist. Z. B. ist ω Limes der Menge $W(\omega)$ aller endlichen Zahlen $0, 1, 2, \ldots, \nu, \ldots$, aber auch jeder unendlichen Menge von endlichen Zahlen $\alpha_0 < \alpha_1 < \alpha_2 < \cdots$, also $\omega = \lim \nu = \lim \alpha_\nu$.

Zwischen Summe und Limes besteht folgende Beziehung. Ist jeder Zahl $\eta < \beta$ eine Ordnungszahl δ_η zugeordnet, so betrachte man die Summen $(\alpha_0 = 0)$

$$\alpha_\eta = \sum_\xi^{W(\eta)} \delta_\xi, \quad \alpha_\beta = \sum_\eta^{W(\beta)} \delta_\eta,$$

---

[1] Die Gesamtheit aller Ordnungszahlen ist also keine Menge (Kap. I, § 1).

[2] Im Sinne von Kap. IV, § 4 sollte lim sup W geschrieben werden; indessen sind bei wohlgeordneten Mengen untere Limites nie vorhanden.

also

$$\alpha_1 = \delta_0, \quad \alpha_2 = \delta_0 + \delta_1, \quad \alpha_3 = \delta_0 + \delta_1 + \delta_2, \ldots$$
$$\alpha_\omega = \delta_0 + \delta_1 + \delta_2 + \ldots, \quad \alpha_{\omega+1} = \delta_0 + \delta_1 + \delta_2 + \ldots + \delta_\omega, \ldots$$

Allgemein ist dann α_β die kleinste Zahl, die $\geqq \alpha_{\eta+1}$ für jedes η. Daß in der Tat $\alpha_\beta \geqq \alpha_{\eta+1}$, ist evident, und daß α_β die kleinste dieser Zahlen ist, folgt aus der in Kap. IV, § 3 gegebenen Formel für die Abschnitte (Anfangsstrecken) einer Summe, wonach jede Zahl $< \alpha_\beta$ von der Form $\alpha_\eta + \delta_\eta'$ ist (δ_η' ein Abschnitt von δ_η), also ein Abschnitt von $\alpha_\eta + \delta_\eta = \alpha_{\eta+1}$ und demgemäß $< \alpha_{\eta+1}$ für ein bestimmtes η.

Ist insbesondere jedes $\delta_\eta > 0$, so ist $\alpha_\xi < \alpha_\eta$ für $\xi < \eta$; ist außerdem β eine Limeszahl, so ist eine Zahl $\geqq \alpha_{\eta+1}$ auch $> \alpha_\eta$ (für jedes η) und umgekehrt, also ist in diesem Fall

$$\alpha_\beta = \lim \alpha_\eta,$$

die ganze Summe Limes der Menge ihrer „Partialsummen". Z. B. gilt für $\beta = \omega$ die Formel

$$\delta_0 + \delta_1 + \delta_2 + \ldots = \lim_\nu (\delta_0 + \delta_1 + \ldots + \delta_\nu),$$

die der Definition der Summe einer konvergenten Reihe in der Analysis ähnlich ist. Beispielsweise erhält man für $\delta_\nu = \omega^\nu$ ($\delta_0 = 1$)

$$\delta_0 + \delta_1 = 1 + \omega = \omega,$$
$$\delta_0 + \delta_1 + \delta_2 = \omega + \omega^2 = \omega(1 + \omega) = \omega^2 \text{ usw.},$$
$$1 + \omega + \omega^2 + \omega^3 + \ldots = \lim_\nu \omega^\nu.$$

Wir stellen hier noch einige für die Rechnung mit Ordnungszahlen wichtige Formeln zusammen.

(A) Ungleichungen. Aus $\alpha < \beta$ folgt:

$$\mu + \alpha < \mu + \beta,$$
$$\alpha + \mu \leqq \beta + \mu,$$
$$\mu\alpha < \mu\beta \text{ für } \mu > 0,$$
$$\alpha\mu \leqq \beta\mu.$$

Denn wenn $\alpha < \beta$, so ist $\beta = \alpha + \gamma$ ($\gamma > 0$) und vice versa, wie wir schon gesehen haben; folglich

$$\mu + \beta = \mu + \alpha + \gamma > \mu + \alpha,$$
$$\mu\beta = \mu\alpha + \mu\gamma > \mu\alpha,$$

letzteres nur, wenn $\mu\gamma > 0$, also $\mu > 0$. Zu einer Ungleichung darf also links ein Summand resp. ein von Null verschiedener Faktor gesetzt werden. Bei Anbringung eines rechten Summanden oder Faktors kann auf Grund der Teilmengeneigenschaft nur auf $\leqq$ geschlossen werden, nicht auf $<$; als Beispiele dienen

$$1 + \omega = 2 + \omega, \quad 1\,\omega = 2\,\omega.$$

Die Umkehrung dieser Aussagen liefert:

Aus $\mu+\alpha<\mu+\beta$ oder $\alpha+\mu<\beta+\mu$ folgt $\alpha<\beta$.
Aus $\mu+\alpha=\mu+\beta$ folgt $\alpha=\beta$.
Aus $\mu\alpha<\mu\beta$ oder $\alpha\mu<\beta\mu$ folgt $\alpha<\beta$.
Aus $\mu\alpha=\mu\beta$ und $\mu>0$ folgt $\alpha=\beta$.

Aus $\alpha+\mu=\beta+\mu$ folgt $\alpha=\beta$ nur für endliches μ (durch wiederholten Schluß von $\alpha+1=\beta+1$ auf $\alpha=\beta$); für unendliches $\mu\geqq\omega$ ist $\alpha+\mu=\alpha+1+\mu=\alpha+2+\mu=\ldots$. Aus $\alpha\mu=\beta\mu$ folgt $\alpha=\beta$ nur, wenn $\mu=\nu+1$ weder 0 noch eine Limeszahl ist; denn nehmen wir $\alpha\leqq\beta$ an, so ist $\alpha\nu+\alpha=\beta\nu+\beta\geqq\alpha\nu+\beta$, also $\alpha\geqq\beta$ und demnach $\alpha=\beta$. Für $\mu=0$ hingegen ist immer $\alpha\mu=\beta\mu$ und für eine Limeszahl $\mu=\omega\pi$ (vgl. S. 109) $\alpha\mu=\alpha 2\mu=\alpha 3\mu=\ldots$.

(B) Subtraktion. Ist $\beta>\alpha$, so kann man, wie wir gesehen haben, $\beta=\alpha+\xi$ setzen, und die Zahl ξ ist dadurch eindeutig definiert (aus $\alpha+\xi=\alpha+\xi'$ folgt $\xi=\xi'$). Diese eindeutige Auflösung der Gleichung

$$\alpha+\xi=\beta$$

für $\beta>\alpha$ bezeichnen wir mit

$$\xi=-\alpha+\beta$$

(nicht $\beta-\alpha$), so daß

$$\alpha+(-\alpha+\beta)=\beta;$$

wir dehnen dies auch auf den Fall $\beta=\alpha$ aus und haben dann $\xi=0$ zu setzen. ξ ist der Typus der Menge $W(\beta)-W(\alpha)$.

Dagegen hat die Gleichung

$$\eta+\alpha=\beta$$

selbst unter der jedenfalls notwendigen Voraussetzung $\beta\geqq\alpha$ nicht immer eine Auflösung η; z. B. ist die Gleichung $\eta+1=\omega$ unlösbar, da ω der Typus einer Menge ohne letztes Element ist. Wenn sie auflösbar ist, so folgt aus dem oben Gesagten, daß sie für endliches α eine einzige, für unendliches α unendlich viele Lösungen hat, nämlich neben η jedenfalls auch $\eta+1$, $\eta+2,\ldots$ Falls die Gleichung lösbar ist, so bezeichnen wir ihre kleinste Lösung mit

$$\eta=\beta-\alpha,$$

so daß

$$(\beta-\alpha)+\alpha=\beta.$$

Insbesondere ist der unmittelbare Vorgänger einer Zahl β, wenn sie einen hat, mit $\beta-1$, der etwaige unmittelbare Vorgänger dieser Zahl mit $\beta-2$ zu bezeichnen usw.

(C) Division. Ist $\alpha>0$, $\beta>0$, so ist jede Zahl $\zeta<\alpha\beta$ in der Form

$$\zeta=\alpha\eta+\xi \quad (\xi<\alpha,\ \eta<\beta)$$

darstellbar. Das folgt aus der in Kap. IV, § 3 angegebenen Darstellung für die Abschnitte (Anfangsstrecken) eines Produkts.

Ist $\alpha > 0$, so ist jede Ordnungszahl ζ in der Form

$$\zeta = \alpha\eta + \xi \quad (\xi < \alpha) \tag{3}$$

darstellbar. Denn ist β irgend eine Zahl $> \zeta$, also $\zeta < \beta = 1 \cdot \beta \leqq \alpha\beta$, so tritt der ebengenannte Fall ein.

Es kommt noch hinzu, daß ξ, η durch α, ζ eindeutig bestimmt sind. Denn wenn

$$\alpha\eta + \xi = \alpha\eta' + \xi' \quad (\xi < \alpha,\ \xi' < \alpha),$$

so muß zunächst $\eta = \eta'$ sein, denn wäre etwa $\eta < \eta'$, so wäre

$$\alpha\eta + \xi < \alpha\eta + \alpha = \alpha(\eta + 1) \leqq \alpha\eta' \leqq \alpha\eta' + \xi'.$$

Aus $\alpha\eta + \xi = \alpha\eta + \xi'$ folgt dann $\xi = \xi'$.

Die Formel (3) ist der entsprechenden bei endlichen Zahlen ganz analog; gewissermaßen ist η der Quotient, ξ der Rest der Division von ζ durch α.

Z. B. ist $(\alpha = 2)$ jede Zahl in einer der beiden Formen 2η, $2\eta + 1$ darstellbar, also „gerade“ oder „ungerade“; oder $(\alpha = \omega)$ jede Zahl ist in der Form $\omega\eta + \nu$ $(\nu = 0, 1, 2, \ldots)$ darstellbar, eine Limeszahl folglich in der Form $\omega\eta$.

Man kann dieses Verfahren fortsetzen und erhält so eine dem Euklidischen Algorithmus oder der Kettenbruchentwicklung für rationale Zahlen analoge Formelkette. Zwei Zahlen α_0, α_1 $(\alpha_1 > 0)$ bestimmen nach dem Obigen zwei Zahlen η_1, α_2 so, daß

$$\alpha_0 = \alpha_1\eta_1 + \alpha_2 \quad (\alpha_1 > \alpha_2).$$

Ist noch $\alpha_2 > 0$, so bestimmen α_1, α_2 wieder zwei Zahlen η_2, α_3 so, daß

$$\alpha_1 = \alpha_2\eta_2 + \alpha_3 \quad (\alpha_2 > \alpha_3).$$

Ist auch noch $\alpha_3 > 0$, so erhält man wieder

$$\alpha_2 = \alpha_3\eta_3 + \alpha_4 \quad (\alpha_3 > \alpha_4) \text{ usw.}$$

Wegen $\alpha_1 > \alpha_2 > \alpha_3 > \ldots$ muß der Prozeß nach einer endlichen Anzahl von Schritten zu Ende kommen, da es keine Menge von Ordnungszahlen gibt, die den Typus ω^* hat (jede Menge von Ordnungszahlen ist wohlgeordnet). Es muß also für eine natürliche Zahl n sich ergeben

$$\alpha_{n-2} = \alpha_{n-1}\eta_{n-1} + \alpha_n \quad (\alpha_{n-1} > \alpha_n > 0),$$
$$\alpha_{n-1} = \alpha_n\eta_n.$$

Beispiel: für $\alpha_0 = \omega 3 + 7$, $\alpha_1 = \omega 2 + 5$ findet man:

$$\begin{aligned} \omega 3 + 7 &= (\omega 2 + 5)\cdot 1 + (\omega + 7), \\ \omega 2 + 5 &= (\omega + 7)\cdot 1 \quad + (\omega + 5), \\ \omega + 7 &= (\omega + 5)\cdot 1 \quad + 2, \\ \omega + 5 &= 2\cdot(\omega + 2) \quad + 1, \\ 2 &= 1\cdot 2. \end{aligned}$$

Wir kommen auf diesen Kettenbruchprozeß später zurück (Kap. VI, § 9).

(D) Limesformeln. Ist $W = \{\alpha_0, \alpha_1, \ldots, \alpha_\eta, \ldots\}$ eine der Größe nach geordnete Menge von Ordnungszahlen ohne letztes Glied, ihr Typus β also eine Limeszahl, so ist

(4) $$\lim_\eta (\mu + \alpha_\eta) = \mu + \lim_\eta \alpha_\eta,$$

(5) $$\lim_\eta \mu\,\alpha_\eta = \mu \cdot \lim_\eta \alpha_\eta \quad (\mu > 0).$$

Beweis. Es sei $\alpha_\beta = \lim \alpha_\eta$, also die erste auf alle Zahlen von W folgende Zahl. Zunächst ist, wenn $\xi < \eta < \beta$, $\alpha_\xi < \alpha_\eta$, also

$$\mu + \alpha_\xi < \mu + \alpha_\eta,$$

$$\mu\,\alpha_\xi < \mu\,\alpha_\eta \quad (\text{für } \mu > 0);$$

die Zahlen $\mu + \alpha_\eta$ und, für $\mu > 0$, die Zahlen $\mu\,\alpha_\eta$ bilden also wieder eine Menge (vom Typus β) von Ordnungszahlen. Ferner ist

$$\mu + \alpha_\beta > \mu + \alpha_\eta;$$

die Zahl $\mu + \alpha_\beta$ ist also größer als alle Zahlen $\mu + \alpha_\eta$. Sie ist aber die kleinste Zahl dieser Eigenschaft, denn soll eine Zahl $> \mu + \alpha_\eta$ sein, für jedes η, also jedenfalls $> \mu$, so können wir sie in der Form $\mu + \alpha$ annehmen, und dann folgt aus $\mu + \alpha > \mu + \alpha_\eta$, daß $\alpha > \alpha_\eta$, also $\alpha \geqq \alpha_\beta$, demnach $\mu + \alpha \geqq \mu + \alpha_\beta$.

Ebenso ist, für $\mu > 0$,

$$\mu\,\alpha_\beta > \mu\,\alpha_\eta.$$

Daß $\mu\,\alpha_\beta$ wieder die kleinste Zahl dieser Art ist, folgt so: jede Zahl ζ kann nach (3) in der Form

$$\zeta = \mu\,\alpha + \nu \quad (\nu < \mu)$$

angenommen werden. Soll $\zeta > \mu\,\alpha_\eta$ sein, so folgt

$$\mu\,\alpha_\eta < \mu\,\alpha + \nu < \mu(\alpha + 1),$$

also $\alpha + 1 > \alpha_\eta$, folglich $\alpha + 1 \geqq \alpha_\beta$. Da α_β Limeszahl ist, so ist hier das Gleichheitszeichen ausgeschlossen, also $\alpha + 1 > \alpha_\beta$, $\alpha \geqq \alpha_\beta$ und

$$\zeta \geqq \mu\,\alpha \geqq \mu\,\alpha_\beta.$$

Damit sind die fraglichen Formeln bewiesen.

Die Nachsetzung des Summanden resp. Faktors μ in diesen

Formeln ist schon aus dem Grunde unzulässig, weil die Zahlen $\alpha_\eta + \mu$ und $\alpha_\eta \mu$ gar nicht mehr eine wohlgeordnete Menge ohne letztes Glied zu bilden brauchen.

Beispiele: durchläuft ν die endlichen Ordnungszahlen $0, 1, 2, \ldots$, so ist

$$\lim \nu = \omega,$$

also

$$\lim (\omega + \nu) = \omega + \omega = \omega 2,$$
$$\lim (\omega 2 + \nu) = \omega 2 + \omega = \omega 3,$$
$$\lim \omega \nu = \omega \omega = \omega^2 \text{ usw.}$$

Wir schließen diesen Paragraphen mit einer Darstellung des Anfangs der Zahlenreihe, d. h. der Menge $W(\alpha)$ für hinlänglich großes α. Zunächst kommen die endlichen Zahlen

$$0, 1, 2, \ldots$$

Auf diese folgt, als Typus der Menge dieser Zahlen, die Zahl ω, der dann $\omega + 1$, $\omega + 2$ usw. folgen, also

$$\omega, \ \omega + 1, \ \omega + 2, \ \ldots$$

Auf diese Zahlen folgt, als Typus der Menge

$$\{0, 1, 2, \ldots, \omega, \omega + 1, \omega + 2, \ldots\},$$

die aus zwei Folgen besteht, die Zahl $\omega + \omega = \omega 2$ und demnach

$$\omega 2, \ \omega 2 + 1, \ \omega 2 + 2, \ldots$$

Auf diese wieder folgt, als Typus einer aus drei Folgen bestehenden Menge, die Zahl $\omega 3$, also

$$\omega 3, \ \omega 3 + 1, \ \omega 3 + 2, \ \ldots \text{ usw.}$$

Die Menge aller dieser Folgen, eine Folge von Folgen, hat den Typus

$$\omega + \omega + \omega + \ldots = \omega \omega = \omega^2,$$

und dies ist die erste auf alle Zahlen $\omega \nu_1 + \nu_0$ (ν_0, ν_1 endlich) folgende Zahl, aber auch die erste auf alle Zahlen $\omega \nu$ folgende Zahl.

Der weitere Fortgang in der Reihe der Ordnungszahlen vollzieht sich immer nach dem Prinzip, daß α der Typus der Menge $W(\alpha)$ aller vorangehenden Zahlen ist. Wir deuten noch die nächsten Schritte an. Auf ω^2 folgen unmittelbar $\omega^2 + 1, \omega^2 + 2, \ldots, \omega^2 + \omega$ usw., also die Zahlen

$$\omega^2 + \omega \nu_1 + \nu_0$$

in derselben Reihenfolge wie vor ω^2 die Zahlen $\omega \nu_1 + \nu_0$. Die Menge aller dieser Zahlen und der vorhergehenden besteht jetzt aus zwei Doppelfolgen vom Typus ω^2, hat also den Typus $\omega^2 + \omega^2 = \omega^2 2$. Es folgen also nunmehr die Zahlen

$$\omega^2 2 + \omega \nu_1 + \nu_0,$$

dann die Zahlen

$$\omega^2 3 + \omega v_1 + v_0 \quad \text{usw.},$$

und auf alle bisherigen Zahlen

$$\omega^2 v_2 + \omega v_1 + v_0 \quad (v_0, v_1, v_2 \text{ endlich})$$

folgt als Limes der Typus einer Folge von Doppelfolgen, also

$$\omega^2 + \omega^2 + \omega^2 + \dots = \omega^2 \omega = \omega^3,$$

der Typus einer dreifachen Folge und zugleich, beispielsweise, die erste auf alle Zahlen $\omega^2 v$ folgende Zahl ($\lim \omega^2 v = \omega^2 \omega = \omega^3$).

Den Anfang von $W(\alpha)$ bei hinreichend großem α machen also die Polynome in ω

$$\omega^m v_m + \omega^{m-1} v_{m-1} + \dots + \omega^2 v_2 + \omega v_1 + v_0$$

mit endlichen Koeffizienten $v_0 v_1 \dots v_m$. Die nächste auf alle diese folgende Zahl, oder der Limes der Menge $\{\omega, \omega^2, \omega^3, \dots\}$ ist in der bisherigen Form nicht ausdrückbar und verlangt, wenn man nicht bei einer Darstellung wie

$$\omega + \omega^2 + \omega^3 + \dots$$

bleiben will, die Einführung eines neuen Zeichens. Durch die in § 4 zu besprechende Definition von Produkten und Potenzen kann diese Einführung an eine spätere Stelle verschoben und für die jetzt in Betracht kommende Zahl noch das Zeichen ω^ω verwandt werden.

§ 3. Transfinite Induktion.

Der Leser wird an dieser Stelle gern einen kurzen Blick rückwärts tun und der genialen Schöpfung G. Cantors, dem System der Ordnungszahlen, seine Bewunderung nicht versagen. Die letzten Betrachtungen über den Anfang dieses Systems zeigen, daß die paradox scheinende Idee, über die endliche Zahlenreihe hinaus den Zählprozeß fortzusetzen, wirklich ausführbar ist, und zwar nicht in einer nebelhaften Weise mit fragwürdigen Unendlichkeitssymbolen wie ∞, sondern nach einem präzisen Gesetz, das an jeder Stelle des Zahlensystems die nunmehr folgende Zahl als Typus der Menge aller vorangehenden Zahlen eindeutig bestimmt. Für die wohlgeordneten Mengen ist damit auch das Postulat erfüllt, daß der Zählprozeß nicht nur die ganze Menge, sondern auch ihre Elemente „zählen", ihnen nämlich bestimmte Zahlzeichen als Nummern oder Indices zuordnen solle; denn die Elemente einer wohlgeordneten Menge A vom Typus α werden eben durch die Zahlen der Menge $W(\alpha)$ in diesem Sinne gezählt, d. h. umkehrbar eindeutig repräsentiert.

Die Tragweite dieser Theorie der Ordnungszahlen werden wir noch an vielen Stellen kennen zu lernen haben. Hier muß zunächst auf eine weitgehende Analogie mit der Reihe der endlichen Zahlen

(als die wir, der formalen Übereinstimmung wegen, die der natürlichen Zahlen einschließlich 0, also die Reihe $\nu = 0, 1, 2, \ldots$ betrachten) hingewiesen werden, nämlich auf die Anwendbarkeit des sogenannten vollständigen Induktionsschlusses. Der Leser kennt aus zahllosen Beispielen den Schluß von ν auf $\nu + 1$, der besagt: eine Aussage $f(\nu)$ bezüglich der endlichen Zahl ν ist für jedes ν richtig, falls $f(0)$ richtig ist und falls aus der Richtigkeit von $f(\nu)$ auf die von $f(\nu + 1)$ geschlossen werden kann. Im Gebiete der endlichen und unendlichen Ordnungszahlen gilt nun ein ähnliches Schlußverfahren, nämlich:[1]

Eine Aussage $f(\alpha)$ bezüglich der Ordnungszahl α ist für jedes α richtig, sobald $f(0)$ richtig ist und sobald aus der Richtigkeit aller $f(\xi)$ für $\xi < \alpha$ auf die Richtigkeit von $f(\alpha)$ geschlossen werden kann.

In der Tat, wäre etwa $f(\beta)$ unrichtig, so ist die Menge derjenigen Zahlen von $W(\beta + 1)$ (d. h. der Zahlen $\leqq \beta$), für die $f(\alpha)$ unrichtig ist, von Null verschieden und hat als Menge von Ordnungszahlen ein erstes Element α; es wäre also α die kleinste Zahl, für die $f(\alpha)$ unrichtig ist. Hier ist $\alpha > 0$, da $f(0)$ als richtig angenommen war. Für alle Zahlen $\xi < \alpha$ ist $f(\xi)$ richtig, daraus sollte aber vorausgesetztermaßen die Richtigkeit von $f(\alpha)$ folgen, und die Annahme, $f(\alpha)$ sei unrichtig, ist also zu verwerfen.

Wir werden diesen vollständigen Induktionsschluß im Gebiete der endlichen und unendlichen Ordnungszahlen (transfinite Induktion) häufig anzuwenden haben.

Wir können aber nicht nur induktiv schließen, sondern auch induktiv definieren. Bedeutet $f(\alpha)$ jetzt nicht eine Aussage hinsichtlich α, sondern ein der Zahl α zugeordnetes Ding, eine Funktion von α (z. B. kann $f(\alpha)$ wieder eine durch α bestimmte Ordnungszahl sein, etwa $f(\alpha) = \alpha^2$, oder eine durch α bestimmte Menge wie etwa $f(\alpha) = W(\alpha)$ u. dgl.), so lautet das induktive Definitionsverfahren:

$f(\alpha)$ ist für jedes α definiert, sobald $f(0)$ definiert ist und sobald vermöge der Definition aller $f(\xi)$ für $\xi < \alpha$ auch $f(\alpha)$ definiert ist.

Es ist evident, wie man den Wortlaut dieser Sätze abzuändern hat, wenn man die Behauptung $f(\alpha)$ oder die Funktion $f(\alpha)$ nur für die Zahlen $\alpha \geqq \beta$ (statt $\alpha \geqq 0$) beweisen resp. definieren will; man hat dann $f(\beta)$, statt $f(0)$, als richtig resp. definiert anzunehmen.

Wir wollen von der transfiniten Induktion eine Anwendung

---

[1] Eine Aussage (z. B. $\alpha + 1 > \alpha$) kann sich auf jede Ordnungszahl beziehen, obwohl die Gesamtheit aller Ordnungszahlen keine widerspruchfreie Menge ist; dasselbe gilt von einer Funktion von α.

machen, von der im folgenden mehrfach besondere Fälle auftreten werden. Jeder Ordnungszahl α sei bereits eine zweite Ordnungszahl $f(\alpha)$ zugeordnet. Die hiermit definierte Funktion $f(\alpha)$ heiße eine Normalfunktion[1], wenn sie monoton wächst, d. h.

$$f(\alpha) < f(\beta) \quad \text{für} \quad \alpha < \beta,$$

und wenn sie der Limesbedingung genügt

$$f(\alpha) = \lim f(\xi) \quad \text{für} \quad \alpha = \lim \xi,$$

wobei ξ die Menge $W(\alpha)$ oder eine Teilmenge, mit der $W(\alpha)$ konfinal ist, zu durchlaufen hat. Nach § 2, (A)(D) sind z. B. $f(\alpha) = \mu + \alpha$ und $f(\alpha) = \mu\alpha$ (letztere für $\mu > 0$) Normalfunktionen.

Auf Grund der Monotonie ist stets

$$f(\alpha) \geqq \alpha;$$

denn gäbe es ein $f(\alpha) < \alpha$, so würde vermöge der Funktion $\xi' = f(\xi)$ die Menge $W(\alpha + 1)$ auf eine Teilmenge von $W(\alpha)$ ähnlich abgebildet werden, im Widerspruch gegen den Satz § 2, II.

Andererseits folgt aber auf Grund der Limeseigenschaft, daß nicht durchweg $f(\alpha) > \alpha$ sein kann, sondern daß es Zahlen α geben muß, und sogar unendlich viele, für die $f(\alpha) = \alpha$. Wir wollen diese Zahlen kritische Zahlen für die Normalfunktion $f(\alpha)$ nennen; diejenigen, wo $f(\alpha) > \alpha$, gewöhnliche Zahlen. Ist nämlich α_0 eine gewöhnliche Zahl, so bilden wir

$$\alpha_1 = f(\alpha_0), \ \alpha_2 = f(\alpha_1), \ \alpha_3 = f(\alpha_2), \ \ldots$$

Es ist dann, weil $f(\alpha)$ monoton wächst und α_0 eine gewöhnliche Zahl ist:

$$\begin{aligned} &\alpha_0 < \alpha_1, \ f(\alpha_0) < f(\alpha_1), \\ &\alpha_1 < \alpha_2, \ f(\alpha_1) < f(\alpha_2), \\ &\alpha_2 < \alpha_3, \ f(\alpha_2) < f(\alpha_3) \quad \text{usw.} \end{aligned}$$

Bilden wir jetzt den Limes

$$\alpha = \lim \alpha_\nu = \lim \alpha_{\nu+1},$$

so folgt auf Grund der Limeseigenschaft der Normalfunktionen

$$f(\alpha) = \lim f(\alpha_\nu) = \lim \alpha_{\nu+1} = \alpha.$$

Demnach ist α eine kritische Zahl, übrigens die kleinste kritische Zahl $> \alpha_0$; denn soll $\beta = f(\beta)$ und $\beta > \alpha_0$ sein, so folgt $f(\beta) > f(\alpha_0)$, also $\beta > \alpha_1$, dann $f(\beta) > f(\alpha_1)$, also $\beta > \alpha_2$ usw., also $\beta \geqq \lim \alpha_\nu$.

Dieses Raisonnement zeigt überdies, daß es über jeder Menge von kritischen Zahlen eine größere kritische Zahl gibt; denn ist K

---

[1] Man könnte sie auch eine monotone stetige Funktion nennen, nach Analogie der Stetigkeitsbedingung $f(\lim x) = \lim f(x)$ für reelle Funktionen einer reellen Variablen.

eine Menge kritischer Zahlen und α_0 eine Zahl, die allen Zahlen von K folgt, so ist α_0 entweder selbst kritisch oder, wenn es eine gewöhnliche Zahl ist, so liefert das obige Verfahren eine kritische Zahl $\alpha > \alpha_0$.

Der Limes einer Menge kritischer Zahlen ist wieder eine kritische Zahl; denn ist $\alpha_\beta = \lim_\eta \alpha_\eta$ $(\eta < \beta)$ und sind die α_η kritische Zahlen, so ist

$$f(\alpha_\beta) = \lim_\eta f(\alpha_\eta) = \lim_\eta \alpha_\eta = \alpha_\beta.$$

Wir definieren nunmehr durch transfinite Induktion eine Funktion $\varkappa_\alpha = g(\alpha)$ auf folgende Weise: $\varkappa_0$ sei die kleinste kritische Zahl für die Normalfunktion $f(\alpha)$, und für $\alpha > 0$ sei $\varkappa_\alpha$ die kleinste kritische Zahl, die auf alle kritischen Zahlen $\varkappa_\xi$ (für $\xi < \alpha$) folgt. Jede kritische Zahl erhält damit einen Index, nämlich den Typus der Menge aller vorhergehenden kritischen Zahlen. Es ist evident, daß die Funktion $g(\alpha)$ wieder eine Normalfunktion ist; ihr Wachstum mit wachsendem Index folgt aus der Definition, und für $\alpha = \lim \xi$ ist $\varkappa_\alpha = \lim \varkappa_\xi$, denn die rechte Seite ist eine kritische Zahl und die kleinste Zahl $> \varkappa_\xi$, also die kleinste kritische Zahl $> \varkappa_\xi$. Also ist wieder $g(\alpha) \geqq \alpha$, es gibt abermals für diese Normalfunktion kritische Zahlen $\alpha = g(\alpha)$, und daraus entspringt eine dritte Normalfunktion $h(\alpha)$ usw.

Beispiele. Ist μ eine feste Ordnungszahl, so ist $f(\alpha) = \mu + \alpha$ eine Normalfunktion. Nehmen wir $\mu > 0$ an (für $\mu = 0$, $f(\alpha) = \alpha$ sind alle Zahlen kritisch), so ist $f(0) > 0$, die Zahl 0 also gewöhnlich. Die kleinste kritische Zahl ist der Limes von

$$f(0) = \mu,\ f(\mu) = \mu 2,\ f(\mu 2) = \mu 3,\ \ldots,$$

d. h. die Zahl $\mu\omega$; in der Tat ist

$$f(\mu\omega) = \mu + \mu\omega = \mu(1 + \omega) = \mu\omega.$$

Von da an sind alle Zahlen kritisch, wie man leicht sieht, also

$$\varkappa_\alpha = g(\alpha) = \mu\omega + \alpha.$$

Die Funktion $f(\alpha) = \alpha + \mu$, $\mu > 0$ vorausgesetzt, gibt stets $f(\alpha) > \alpha$; woraus zu schließen ist, daß sie keine Normalfunktion ist. In der Tat braucht $f(\alpha)$ weder monoton zu wachsen noch ist die Limeseigenschaft vorhanden (z. B. für $\mu = 1$ ist $f(\omega) = \omega + 1$ nicht der Limes der Zahlen $f(0), f(1), f(2), \ldots = 1, 2, 3, \ldots$).

Ist $\mu > 0$, so ist auch $f(\alpha) = \mu\alpha$ eine Normalfunktion. Lassen wir den trivalen Fall $\mu = 1$ beiseite, wo alle Zahlen kritisch sind, und nehmen $\mu > 1$ an, so ist die Zahl 0 kritisch, weil $f(0) = 0$; die

Zahl 1 ist gewöhnlich, weil $f(1) = \mu > 1$, und die erste kritische Zahl > 1 erhält man also als Limes[1] der Folge

$$1,\ f(1) = \mu,\ f(\mu) = \mu^2,\ f(\mu^2) = \mu^3,\ \ldots$$

Nennen wir diese erste von 0 verschiedene kritische Zahl $\varkappa$ (also $\varkappa_0 = 0$, $\varkappa_1 = \varkappa$), so können wir alle kritischen Zahlen leicht bestimmen. Nach § 2 (C) ist jede Zahl α in der Form

$$\alpha = \varkappa\eta + \xi \qquad (\xi < \varkappa)$$

darstellbar. Daraus folgt

$$f(\alpha) = \mu(\varkappa\eta + \xi) = \mu\varkappa\eta + \mu\xi = \varkappa\eta + \mu\xi.$$

Nun ist für $0 < \xi < \varkappa$ die Zahl ξ gewöhnlich, also $\mu\xi > \xi$, folglich

$$f(\alpha) > \varkappa\eta + \xi = \alpha.$$

Also kann α nur für $\xi = 0$ kritisch sein, d. h. die kritischen Zahlen müssen von der Form $\varkappa\eta$ sein, welche Bedingung auch hinreicht. Also ist

$$\varkappa_\alpha = g(\alpha) = \varkappa\alpha,$$

die kritischen Zahlen sind $0, \varkappa, \varkappa 2, \varkappa 3, \ldots$

Die Funktion $f(\alpha) = \alpha\mu$, $\mu > 1$, ist keine Normalfunktion, da für $\alpha > 0$ stets $f(\alpha) > \alpha$ ist.

Die Funktion $f(\alpha) = \alpha^2$ ist ebenfalls keine, da für $\alpha > 1$ stets $f(\alpha) > \alpha$ ist.

Allgemeinere Normalfunktionen erhalten wir durch Iteration einer Funktion, die selbst keine Normalfunktion ist. Sei jeder Ordnungszahl α eine zweite Ordnungszahl $\varphi(\alpha) > \alpha$ zugeordnet. Wir definieren dann durch Induktion die Funktion $f(\alpha)$ folgendermaßen: $f(0)$ sei eine beliebig gegebene Ordnungszahl, und für $\alpha > 0$ sei $f(\alpha)$ die kleinste Zahl, die $\geqq \varphi(f(\xi))$ für jedes $\xi < \alpha$.

Für $\alpha < \beta$ hat man

$$\varphi(f(\beta)) > f(\beta) \geqq \varphi(f(\alpha)) > f(\alpha);$$

$f(\alpha)$ und $\varphi(f(\alpha))$ sind also monoton wachsende Funktionen. Danach ist für eine Zahl $\alpha + 1$ mit unmittelbarem Vorgänger

$$f(\alpha + 1) = \varphi(f(\alpha));$$

denn für $\xi < \alpha + 1$ (d. h. $\xi \leqq \alpha$) ist $\varphi(f(\xi)) \leqq \varphi(f(\alpha))$. Dagegen ist für eine Limeszahl α

$$f(\alpha) = \lim_{\xi} f(\xi) \qquad (\xi < \alpha),$$

$f(\alpha)$ also eine Normalfunktion. Denn nach Definition soll

$$f(\alpha) \geqq \varphi(f(\xi)) = f(\xi + 1) > f(\xi)$$

---

[1] Unter Vorwegnahme der Potenzdefinition des nächsten Paragraphen würde dieser Limes mit μ^ω zu bezeichnen sein, wonach sich tatsächlich ergibt

$$f(\mu^\omega) = \mu \cdot \mu^\omega = \mu^{1+\omega} = \mu^\omega.$$

sein, $f(\alpha)$ muß also alle $f(\xi)$ übertreffen. Umgekehrt, eine Zahl, die alle $f(\xi)$ übertrifft, übertrifft auch alle $f(\xi+1)$ (da $\xi+1$ mit ξ zu $W(\alpha)$ gehört), ist also sicher $\geqq \varphi(f(\xi))$. Unter diesen Zahlen ist $\lim f(\xi)$ die kleinste, also stimmt $f(\alpha)$ mit diesem Limes überein.

Den Anfang der $f(\alpha)$ bilden, wenn wir kürzer $f(\alpha)=\beta_\alpha$ schreiben, die Zahlen

$$\beta_0,\ \beta_1=\varphi(\beta_0),\ \beta_2=\varphi(\beta_1),\ \beta_3=\varphi(\beta_2),\ \ldots$$
$$\beta_\omega=\lim\beta_\nu,\ \beta_{\omega+1}=\varphi(\beta_\omega),\ \ldots$$

Wenn die erzeugende Funktion $\varphi(\alpha)$ die Ungleichung $\varphi(\alpha)>\alpha$ nicht durchweg, sondern erst für $\alpha\geqq\alpha_0$ erfüllt, so muß man auch $f(0)\geqq\alpha_0$ wählen; im übrigen bleibt alles unverändert.

Wählen wir z. B. als erzeugende Funktion $\varphi(\alpha)=\alpha+\mu$ für $\mu>0$ und $f(0)=0$, so wird

$$\beta_0=0,\ \beta_1=\mu,\ \beta_2=\mu 2,\ \ldots,\ \beta_\omega=\mu\omega,\ \beta_{\omega+1}=\mu(\omega+1),\ \ldots$$

allgemein $f(\alpha)=\mu\alpha$; Multiplikation ist hier also durch iterierte Addition gewonnen. Dies zeigt den Weg, Potenzierung durch iterierte Multiplikation zu definieren.

§ 4. Potenzen und Produkte.

Wir haben, für den Fall ungeordneter Mengen, Potenzen und Produkte als Mengen von Elementkomplexen definiert; für den Fall geordneter Mengen beschränkten wir uns, um diese Mengen von Elementkomplexen lexikographisch anordnen zu können, auf den Fall von Produkten aus endlich vielen Faktoren. Im nächsten Kapitel werden wir auch geordnete Produkte unendlich vieler Faktoren definieren, als lexikographisch geordnete Teilmengen jener Komplexmengen (die sich als ganzes im allgemeinen nicht lexikographisch ordnen lassen). Anscheinend ganz selbständig und isoliert steht die in diesem Paragraphen zu gebende Erklärung von Produkten und Potenzen von Ordnungszahlen; im nächsten Kapitel werden wir sie aber als Spezialfall in das allgemeine System einreihen. Wir haben diese Bemerkung vorausgeschickt, um den Leser von vornherein vor einem Fehlschluß bezüglich der Mächtigkeiten zu warnen: während nach unseren Definitionen der Typus $\alpha+\beta$ die Mächtigkeit $\mathfrak{a}+\mathfrak{b}$, der Typus $\alpha\beta$ die Mächtigkeit $\mathfrak{a}\mathfrak{b}$ hat (falls α,β die Mächtigkeiten $\mathfrak{a},\mathfrak{b}$ haben), wird die Potenz α^β, die wir jetzt definieren, keineswegs die Mächtigkeit $\mathfrak{a}^\mathfrak{b}$ haben — was sich auf dem jetzigen Standpunkt so erklärt, daß wir augenblicklich mit jenen Komplexmengen überhaupt nicht operieren, und im nächsten Kapitel dahin erklären wird, daß die Menge vom Typus α^β nur ein Teil jener Komplexmenge von der Mächtigkeit $\mathfrak{a}^\mathfrak{b}$ ist. Z. B. werden wir $2^\omega=\omega$ haben, während $2^{\aleph_0}>\aleph_0$ ist.

Wir definieren zunächst die **Potenz** μ^α, indem wir sie bei fester Basis μ als Funktion des Exponenten α ansehen, durch Induktion, indem wir μ^α durch die sämtlichen μ^ξ für $\xi < \alpha$ ausdrücken. Mit Ausschließung trivialer Fälle nehmen wir die Basis > 1 an.[1] Wir definieren dann $\mu^0 = 1$ und, für $\alpha > 0$, μ^α als die kleinste Zahl $\geqq \mu^\xi \cdot \mu$.

Dies ist also (vgl. den Schluß von § 3) ein Spezialfall der Erzeugung einer Normalfunktion $f(\alpha)$ durch Iteration einer Funktion $\varphi(\alpha) > \alpha$. Setzen wir nämlich $\varphi(\alpha) = \alpha\mu$, mit $\mu > 1$, so ist $\varphi(\alpha) > \alpha$ wenigstens für $\alpha \geqq 1$; definieren wir also $f(0) = 1$ und $f(\alpha)$ als kleinste Zahl $\geqq f(\xi)\cdot\mu$, so stimmt $f(\alpha)$ mit dem soeben erklärten μ^α überein. Die Potenzierung wird durch iterierte Multiplikation erklärt.

Da μ^α eine Normalfunktion ist, so haben wir

(1) $$\mu^\alpha < \mu^\beta \text{ für } \alpha < \beta,$$

(2) $$\mu^\alpha = \lim \mu^\xi \text{ für } \alpha = \lim \xi,$$

wobei ξ die Menge $W(\alpha)$ oder eine Teilmenge, mit der $W(\alpha)$ konfinal ist, zu durchlaufen hat; wir schreiben ausführlicher

(3) $$\mu^{\alpha_\beta} = \lim_\eta \mu^{\alpha_\eta} \text{ für } \alpha_\beta = \lim_\eta \alpha_\eta.$$

Endlich ist für eine Zahl mit Vorgänger

(4) $$\mu^{\alpha+1} = \mu^\alpha \cdot \mu.$$

Beispiele. Es ist stets

$$\mu^1 = \mu^0 \cdot \mu = \mu, \quad \mu^2 = \mu^1 \cdot \mu = \mu\cdot\mu, \quad \mu^3 = \mu^2\cdot\mu = \mu\cdot\mu\cdot\mu, \ \ldots,$$

wovon wir ja schon Gebrauch gemacht haben.

μ^ω ist der Limes über der Menge dieser Zahlen, z. B.

$$2^\omega = \lim 2^\nu = \omega, \quad \text{ebenso } \omega = 3^\omega = 4^\omega = \ldots;$$

$$\omega^\omega = \lim \omega^\nu = \omega + \omega^2 + \omega^3 + \ldots$$

Weiter ist $\mu^{\omega+1} = \mu^\omega \cdot \mu$, z. B.

$$2^{\omega+1} = 2^\omega \cdot 2 = \omega \cdot 2 = \omega + \omega,$$

$$\omega^{\omega+1} = \omega^\omega \cdot \omega = \omega^\omega + \omega^\omega + \omega^\omega + \ldots$$

Es gelten die **Potenzregeln**

(5) $$\mu^\alpha \cdot \mu^\beta = \mu^{\alpha+\beta}, \quad (\mu^\alpha)^\beta = \mu^{\alpha\beta},$$

die sich als Spezialfälle des assoziativen Gesetzes der Multiplikation auffassen lassen; da das kommutative Gesetz nicht gilt, so ist die bei endlichen Zahlen giltige Regel

$$\mu^\alpha \cdot \nu^\alpha = (\mu\nu)^\alpha$$

---

[1] 1^α wäre $= 1$, 0^α für $\alpha > 0$ jedenfalls $= 0$ zu definieren; bezüglich 0^0 scheint jeder der beiden Werte 0 und 1 zulässig. Diese Definitionen haben natürlich nur den Zweck, die unten folgenden Potenzformeln allgemeingültig zu machen.

hier nicht richtig (z. B. sind schon $\mu^2\nu^2 = \mu\mu\nu\nu$ und $(\mu\nu)^2 = \mu\nu\mu\nu$ im allgemeinen verschieden).

Wir beweisen die Potenzregeln wieder durch Induktion, indem wir zeigen, daß sie für $\beta > 0$ richtig sind, falls sie für jedes $\eta < \beta$ richtig sind (für $\beta = 0$ sind sie richtig). Statt dabei auf die allgemeine Definition der Potenz zurückzugreifen, können wir auch die Formeln (2), (3) und (4) benutzen.

Hat $\beta = \gamma + 1$ einen unmittelbaren Vorgänger $\gamma = \beta - 1$, so ist

$$\mu^\alpha \cdot \mu^\beta = \mu^\alpha \cdot \mu^{\gamma+1} = \mu^\alpha \cdot \mu^\gamma \cdot \mu = \mu^{\alpha+\gamma} \cdot \mu = \mu^{\alpha+\gamma+1} = \mu^{\alpha+\beta},$$

$$(\mu^\alpha)^\beta = (\mu^\alpha)^{\gamma+1} = (\mu^\alpha)^\gamma \cdot \mu^\alpha = \mu^{\alpha\gamma} \cdot \mu^\alpha = \mu^{\alpha\gamma+\alpha} = \mu^{\alpha(\gamma+1)} = \mu^{\alpha\beta};$$

bei dem Beweis der zweiten Formel wird schon die erste benutzt.

Ist β Limeszahl, $\beta = \lim \eta$, so ist

$$\mu^\alpha \cdot \mu^\beta = \mu^\alpha \cdot \lim \mu^\eta = \lim \mu^\alpha \cdot \mu^\eta = \lim \mu^{\alpha+\eta}$$
$$= \mu^{\lim(\alpha+\eta)} = \mu^{\alpha+\lim \eta} = \mu^{\alpha+\beta};$$

und für $\alpha > 0$ (die Potenzregeln sind für $\alpha = 0$ trivial)

$$(\mu^\alpha)^\beta = \lim (\mu^\alpha)^\eta = \lim \mu^{\alpha\eta} = \mu^{\lim \alpha\eta} = \mu^{\alpha \cdot \lim \eta} = \mu^{\alpha\beta}.$$

Um ein **Produkt** von Ordnungszahlen zu definieren, denken wir uns jeder Ordnungszahl α wieder eine Ordnungszahl μ_α zugeordnet und sehen das zu definierende Produkt

$$\prod_\xi^{W(\alpha)} \mu_\xi = \mu_0 \mu_1 \dots \mu_\xi \dots$$

als eine Funktion von α an, welche Zahl wir auch den Exponenten des Produkts nennen (nicht das Argument). Diese Funktion $f(\alpha)$, für die ersten Zahlen also

$$f(1) = \mu_0, \quad f(2) = \mu_0 \mu_1, \quad f(3) = \mu_0 \mu_1 \mu_2, \ \dots$$

definieren wir durch Induktion als erste Zahl $\geqq f(\xi) \cdot \mu_\xi$ für alle $\xi < \alpha$. Der Leser wird sich leicht überzeugen, daß, analog zu (2) und (4),

$$f(\alpha + 1) = f(\alpha) \cdot \mu_\alpha$$

und, falls alle Faktoren > 1 sind, für eine Limeszahl α

$$f(\alpha) = \lim_\xi f(\xi)$$

ist. Das assoziative Gesetz für diese Multiplikation zu entwickeln verzichten wir, da wir es später allgemein zu beweisen haben werden (Kap. VI, § 4).

Beispiele:

$$2 \cdot 2 \cdot 2 \dots = \lim 2^\nu = 2^\omega = \omega,$$
$$1 \cdot 2 \cdot 3 \cdot 4 \dots = \lim \nu! = \omega,$$
$$\omega \cdot \omega \cdot \omega \dots = \lim \omega^\nu = \omega^\omega,$$
$$\omega\omega\omega \dots \omega\omega\omega \dots = \omega^{\omega+\omega} = \omega^{\omega 2} = \omega^\omega \cdot \omega^\omega = (\omega^\omega)^2,$$
$$\omega^\omega \cdot \omega^\omega \cdot \omega^\omega \dots = (\omega^\omega)^\omega = \omega^{\omega\omega} = \omega^{\omega^2} \quad \text{usw.}$$

Mit Hilfe der Potenzen einer Basis $\beta > 1$ läßt sich jede Ordnungszahl in ähnlicher Weise ausdrücken wie etwa bei der üblichen dezimalen Darstellung jede endliche Zahl durch die Potenzen von 10. Unter den Potenzen von β sei β^α die erste, die $> \zeta$ (solche gibt es, da $f(\alpha) = \beta^\alpha$ eine Normalfunktion, also $\beta^\alpha \geqq \alpha$ und demnach z. B. $\beta^{\zeta+1} \geqq \zeta + 1 > \zeta$ ist). Hierbei kann α keine Limeszahl sein, sonst wäre für $\xi < \alpha$ auch $\xi + 1 < \alpha$, also

$$\beta^\xi < \beta^{\xi+1} \leqq \zeta < \beta^\alpha,$$

ζ also eine Zahl zwischen β^ξ und β^α (für jedes ξ), während doch $\beta^\alpha = \lim \beta^\xi$ die erste Zahl nach allen β^ξ ist. Also ist, wenn wir den Fall $\alpha = 0$, $\zeta = 0$ ausschließen, α eine Zahl mit unmittelbarem Vorgänger. Ersetzen wir sie durch $\alpha + 1$, so ist demgemäß durch $\zeta > 0$ eine Zahl α eindeutig bestimmt derart, daß

$$\beta^\alpha \leqq \zeta < \beta^{\alpha+1}.$$

Als Zahl $< \beta^\alpha \cdot \beta$ läßt sich ζ (§ 2, (C)) in der Gestalt

$$\zeta = \beta^\alpha \eta + \zeta_1 \qquad (\eta < \beta,\ \zeta_1 < \beta^\alpha)$$

ausdrücken, wobei η und ζ_1 ebenfalls durch ζ eindeutig bestimmt sind. Ist $\zeta_1 = 0$, so ist die Entwicklung beendet; ist $\zeta_1 > 0$, so bestimmt ζ_1 wieder drei Zahlen $\alpha_1, \eta_1, \zeta_2$ derart, daß

$$\beta^{\alpha_1} \leqq \zeta_1 < \beta^{\alpha_1+1},$$
$$\zeta_1 = \beta^{\alpha_1} \eta_1 + \zeta_2 \quad (\eta_1 < \beta,\ \zeta_2 < \beta^{\alpha_1});$$

ist noch $\zeta_2 > 0$, so erhält man weiter

$$\beta^{\alpha_2} \leqq \zeta_2 < \beta^{\alpha_2+1},$$
$$\zeta_2 = \beta^{\alpha_2} \eta_2 + \zeta_3 \quad (\eta_2 < \beta,\ \zeta_3 < \beta^{\alpha_2}) \text{ usw.}$$

Dabei ist

$$\zeta \geqq \beta^\alpha > \zeta_1 \geqq \beta^{\alpha_1} > \zeta_2 \geqq \beta^{\alpha_2} > \cdots,$$

also

$$\zeta > \zeta_1 > \zeta_2 > \cdots,$$
$$\alpha > \alpha_1 > \alpha_2 > \cdots$$

und $\eta, \eta_1, \eta_2, \ldots$ sind > 0. Da es keine Ordnungszahlenmenge vom Typus ω^* gibt, so muß das Verfahren nach einer endlichen Reihe von Schritten zu einem letzten von 0 verschiedenen

$$\zeta_n = \beta^{\alpha_n} \eta_n$$

führen, und wir haben damit für ζ die Darstellung gewonnen

$$\zeta = \beta^\alpha \eta + \beta^{\alpha_1} \eta_1 + \beta^{\alpha_2} \eta_2 + \cdots + \beta^{\alpha_n} \eta_n,$$

wobei

$$\alpha > \alpha_1 > \alpha_2 > \cdots > \alpha_n (\geqq 0)$$

und

$$0 < \eta, \eta_1, \eta_2, \ldots, \eta_n < \beta;$$

die Koeffizienten η wie die Exponenten α sind durch ζ eindeutig bestimmt. So ist z. B. mit der Basis $\beta = 2$ jede Zahl > 0 in der Gestalt

$$\zeta = 2^{\alpha} + 2^{\alpha_1} + 2^{\alpha_2} + \ldots + 2^{\alpha_n},$$

mit der Basis ω in der Gestalt

$$\zeta = \omega^{\alpha} \nu + \omega^{\alpha_1} \nu_1 + \omega^{\alpha_2} \nu_2 + \ldots + \omega^{\alpha_n} \nu_n$$

mit natürlichen Koeffizienten $(= 1, 2, 3, \ldots)$ darstellbar.

Indessen kann der Fall eintreten, daß man bei dieser Darstellung die Zahl ζ gar nicht durch kleinere Zahlen ausdrücken kann, sondern daß der erste Exponent wieder ζ selbst wird: $\zeta = \beta^{\zeta}$. Das geschieht, wenn ζ eine kritische Zahl für die Normalfunktion $f(\alpha) = \beta^{\alpha}$ ist (immer $\beta > 1$ vorausgesetzt). Da $f(0) = 1 > 0$ ist, für jedes β, so ist die Zahl 0 immer gewöhnlich, und die erste kritische Zahl ist der Limes der Folge

$$f(0) = 1,\ f(1) = \beta,\ f(\beta) = \beta^{\beta},\ f(\beta^{\beta}) = \beta^{\beta^{\beta}}, \ldots$$

Für $\beta = 2$ ist die erste kritische Zahl demnach $\omega = 2^{\omega}$; ebenso für $\beta = 3, 4, \ldots$. Für $\beta = \omega$ ist sie der Limes von

$$1,\ \omega,\ \omega^{\omega},\ \omega^{\omega^{\omega}}, \ldots$$

und für diese Zahl $\varkappa$ würde $\varkappa = \omega^{\varkappa}$ die Darstellung mit der Basis ω sein, wobei also der Exponent wieder $\varkappa$ selbst ist.[1] Diese Zahl macht als erste nach ω die Einführung eines neuen Zeichens notwendig, während sich alle früheren mit Hilfe der Verknüpfungsgesetze, Potenzierung eingeschlossen, durch ω ausdrücken lassen; wir sahen früher, daß ohne Potenzierung bereits ω^{ω} ein neues Zeichen verlangen würde. Im übrigen wissen wir, daß durch Bildung endlicher Komplexe von Zeichen eines endlichen oder abzählbaren Zeichensystems immer nur eine abzählbare Menge von Dingen bezeichnet werden kann, selbst wenn man mehrere Alphabete, Ziffern auf, über und unter der Zeile (Exponenten und Indices), Punkte, Kommata, Klammern, Striche, Spatien usw. als Zeichen zuläßt; hat man also mit einem gegebenen Zeichensystem alle bezeichenbaren Ordnungszahlen gebildet, so verlangt die erste nichtbezeichnete, z. B. größere Ordnungszahl die Einführung eines neuen Zeichens.

Man beachte, daß wir hier nur von endlichen Komplexen der Zeichen des Systems gesprochen haben. Sobald wir Folgen und deren Limeselemente, Summen unendlich vieler Summanden u. dgl. einführen, braucht die Menge der bezeichneten Dinge nicht mehr abzählbar zu sein; die Tatsache aber, daß im Fall einer so bezeichneten Zahlenmenge die nächstgrößere Ordnungszahl ein neues Zeichen fordert, bleibt natürlich bestehen.

---

[1] Diese kritischen Zahlen für die Normalfunktion $f(\alpha) = \omega^{\alpha}$ nennt G. Cantor Epsilonzahlen.

§ 5. Alefs und Zahlenklassen.

Unter einer Kardinalzahl wollen wir jetzt, vorübergehend, nur die Mächtigkeit einer wohlgeordneten Menge verstehen.[1] Zu den Kardinalzahlen gehören die natürlichen Zahlen (und Null) als Mächtigkeiten endlicher Mengen; die Mächtigkeiten unendlicher wohlgeordneter Mengen werden nach dem ersten Buchstaben des hebräischen Alphabets als Alefs bezeichnet. Ein solches Alef ist uns bereits bekannt, nämlich die Mächtigkeit $\aleph_0$ der Menge der natürlichen Zahlen, und wir wissen, daß dies die kleinste unendliche Mächtigkeit überhaupt, um so mehr also das kleinste Alef ist.

Es seien A, B zwei wohlgeordnete Mengen mit den Typen α, β und den Mächtigkeiten $\mathfrak{a}, \mathfrak{b}$. Auf Grund der Vergleichbarkeit der Ordnungszahlen sind drei Fälle möglich:

entweder ist $\alpha = \beta$, A mit B ähnlich; dann ist auch A mit B äquivalent, also $\mathfrak{a} = \mathfrak{b}$.

oder es ist $\alpha < \beta$, A einem Abschnitt von B ähnlich, also A einer Teilmenge von B äquivalent, folglich $\mathfrak{a} \leqq \mathfrak{b}$.

oder es ist $\alpha > \beta$, B einem Abschnitt von A ähnlich, B einer Teilmenge von A äquivalent, folglich $\mathfrak{a} \geqq \mathfrak{b}$.

In jedem Falle sind also die beiden Kardinalzahlen $\mathfrak{a}$ und $\mathfrak{b}$ vergleichbar (Kap. III, § 2). Wir haben ja die Vergleichbarkeit irgend zweier beliebiger Mächtigkeiten bisher nicht beweisen können; jetzt ist uns der Weg dazu gezeigt: sobald wir (§ 7) bewiesen haben, daß jede Menge wohlgeordnet werden kann, daß also jede Mächtigkeit eine Kardinalzahl ist, wird diese Lücke ausgefüllt sein.

Die Zusammenstellung:

$$\text{aus} \quad \alpha < \beta, \quad \alpha = \beta, \quad \alpha > \beta$$
$$\text{folgt} \quad \mathfrak{a} \leqq \mathfrak{b}, \quad \mathfrak{a} = \mathfrak{b}, \quad \mathfrak{a} \geqq \mathfrak{b}$$

liefert umgekehrt:

$$\text{aus} \quad \mathfrak{a} < \mathfrak{b}, \quad \mathfrak{a} > \mathfrak{b}$$
$$\text{folgt} \quad \alpha < \beta, \quad \alpha > \beta,$$

während aus $\mathfrak{a} = \mathfrak{b}$ nichts folgt, d. h. in diesem Falle immer noch jede der drei Relationen $\alpha \lesseqqgtr \beta$ möglich ist. Z. B. kennen wir bereits eine große Menge verschiedener Ordnungszahlen

$$\omega, \ \omega + 1, \ \omega + 2, \ \ldots, \ \omega 2, \ \omega 3, \ \ldots, \ \omega^2, \ \omega^3, \ \ldots,$$

die allesamt die gleiche Mächtigkeit $\aleph_0$ haben.

Die Menge aller verschiedenen Ordnungszahlen von einer gegebenen Mächtigkeit $\mathfrak{a}$ bezeichnen wir mit $Z(\mathfrak{a})$ und nennen sie eine

---

[1] Wir haben früher Kardinalzahl und Mächtigkeit als Synonyma behandelt, und der Unterschied, den wir augenblicklich machen, wird nach dem Beweis des Wohlordnungssatzes (§ 7) auch wieder verschwinden.

Zahlenklasse; es ist dies also eine Teilmenge der Typenklasse $T(\mathfrak{a})$, die ihrerseits alle verschiedenen Ordnungstypen dieser Mächtigkeit $\mathfrak{a}$ umfaßt (Kap. IV, § 7). Z. B. gehören die soeben genannten Ordnungszahlen der Zahlenklasse $Z(\aleph_0)$ an, während die Typenklasse $T(\aleph_0)$ auch noch die Typen nicht wohlgeordneter Mengen, z. B. ω^* und η enthält. Sollte etwa (was aber nachher ausgeschlossen werden wird) $\mathfrak{a}$ keine Kardinalzahl, also die Mächtigkeit einer Menge sein, die auf keine Weise wohlgeordnet werden kann, so wäre $Z(\mathfrak{a})$ die Nullmenge.

Jede Menge K von Kardinalzahlen ist, der Größe nach geordnet, eine wohlgeordnete Menge. Denn sind $\mathfrak{a}, \mathfrak{b}, \ldots$ die Elemente von K, ferner $A, B, \ldots$ irgendwelche wohlgeordnete Mengen dieser Mächtigkeiten und $\alpha, \beta, \ldots$ deren Typen, so folgt aus $\mathfrak{a} \lesseqgtr \mathfrak{b}$ auch $\alpha \lesseqgtr \beta$; die Menge K ist also der Menge der Ordnungszahlen $\alpha, \beta, \ldots$ ähnlich und nach § 2 wohlgeordnet. Insbesondere gibt es in K eine kleinste Kardinalzahl.

Zu jeder Kardinalzahl $\mathfrak{a}$ gibt es eine größere, insbesondere eine nächstgrößere. Bei dem Beweis dieser Behauptung dürfen wir nicht einfach auf das frühere Resultat $2^{\mathfrak{a}} > \mathfrak{a}$ verweisen, solange wir noch nicht wissen, daß auch $2^{\mathfrak{a}}$ eine Kardinalzahl ist. Wir schließen daher so: es sei β die nächstgrößere Ordnungszahl (S. 106) über der Zahlenklasse $Z(\mathfrak{a})$; $\mathfrak{b}$ sei die Mächtigkeit von β, und α irgend eine Ordnungszahl von der Mächtigkeit $\mathfrak{a}$. Aus $\beta > \alpha$ folgt dann $\mathfrak{b} \geqq \mathfrak{a}$; da aber nicht $\mathfrak{b} = \mathfrak{a}$ sein kann, weil sonst β noch zu $Z(\mathfrak{a})$ gehören würde, so ist $\mathfrak{b} > \mathfrak{a}$. Damit ist $\mathfrak{b}$ als eine Kardinalzahl von der gewünschten Eigenschaft erwiesen; überdies ist es die kleinste. Denn soll $\mathfrak{c}$ Kardinalzahl einer wohlgeordneten Menge (vom Typus γ) und $> \mathfrak{a}$ sein, so folgt $\gamma > \alpha$ für jedes α, also $\gamma \geqq \beta$ und demnach $\mathfrak{c} \geqq \mathfrak{b}$.

Über jeder Menge K von Kardinalzahlen gibt es eine größere, insbesondere eine nächstgrößere. Wir können K, als der Größe nach wohlgeordnete Menge vom Typus β, in der Gestalt

$$K = \{\mathfrak{a}_0, \mathfrak{a}_1, \ldots, \mathfrak{a}_\eta, \ldots\} \qquad (\eta < \beta)$$

schreiben. Gibt es in K eine größte Kardinalzahl $\mathfrak{a}_{\beta-1}$, so ist die nächste hierauf folgende Kardinalzahl auch die kleinste Kardinalzahl über K. Hat K kein letztes Element und ist α_η eine Ordnungszahl von der Mächtigkeit $\mathfrak{a}_\eta$, so bilden die Ordnungszahlen α_η ihrerseits eine Menge vom Typus β. Der Limes dieser Zahlen sei

$$\alpha_\beta = \lim_\eta \alpha_\eta$$

und $\mathfrak{a}_\beta$ die zugehörige Mächtigkeit. Dann ist, für jedes $\eta < \beta$,

$$\alpha_\beta > \alpha_{\eta+1}, \ \mathfrak{a}_\beta \geqq \mathfrak{a}_{\eta+1} > \mathfrak{a}_\eta,$$

also $\mathfrak{a}_\beta$ wirklich größer als alle Elemente von K. Zugleich ist $\mathfrak{a}_\beta$ die

kleinste Kardinalzahl dieser Art, denn soll $\mathfrak{a}$ Kardinalzahl einer wohlgeordneten Menge vom Typus α und $\mathfrak{a} > \mathfrak{a}_\eta$ sein für jedes η, so ist auch $\alpha > \alpha_\eta$, also $\alpha \geqq \alpha_\beta$ und $\mathfrak{a} \geqq \mathfrak{a}_\beta$. Wir schreiben auch hier, wie bei den Ordnungszahlen,

$$\mathfrak{a}_\beta = \lim \mathfrak{a}_\eta.$$

Wir weisen nunmehr jeder Ordnungszahl ein Alef zu, d. h. definieren durch transfinite Induktion eine Funktion $f(\alpha) = \aleph_\alpha$ folgendermaßen: $\aleph_0$ sei das uns längst bekannte kleinste Alef, die Mächtigkeit der abzählbaren Mengen, und, für $\alpha > 0$, $\aleph_\alpha$ das kleinste Alef über allen $\aleph_\xi$ $(\xi < \alpha)$. Jedes Alef erhält auf diese Weise einen Index, nämlich den Typus der Menge aller vorangehenden Alefs. Es ist

$$\aleph_\alpha < \aleph_\beta \quad \text{für } \alpha < \beta$$

und

$$\aleph_\alpha = \lim \aleph_\xi \quad \text{für } \alpha = \lim \xi,$$

wobei ξ die Menge $W(\alpha)$ oder eine Teilmenge, mit der $W(\alpha)$ konfinal ist, zu durchlaufen hat. Die Funktion $\aleph_\alpha$ hat also die beiden analogen Eigenschaften wie eine Normalfunktion. Auf das erste Alef $\aleph_0$ folgt als nächstgrößeres $\aleph_1$, dann als nächstgrößeres $\aleph_2$ usw., auf alle Alefs mit endlichen Indices als nächstgrößeres $\aleph_\omega = \lim \aleph_\nu$, dann $\aleph_{\omega+1}$, $\aleph_{\omega+2}$, ..., $\aleph_{\omega 2}$ usw.

Jede Mächtigkeit, die kleiner als eine Kardinalzahl ist, ist selbst eine Kardinalzahl; denn eine wohlgeordnete Menge erteilt auch allen ihren Teilmengen eine Wohlordnung. Auf Grund dieser Bemerkung können wir sagen, daß $\aleph_\alpha$ nicht nur das auf alle $\aleph_\xi$ $(\xi < \alpha)$ nächstfolgende Alef, sondern eine auf alle $\aleph_\xi$ nächstfolgende Mächtigkeit ist in dem Sinne, daß es keine Mächtigkeit zwischen allen $\aleph_\xi$ und $\aleph_\alpha$ gibt; denn eine solche Mächtigkeit wäre ja wieder ein Alef. Solange wir die Vergleichbarkeit aller Mächtigkeiten noch nicht bewiesen haben, dürfen wir allerdings nicht $\aleph_\alpha$ die auf alle $\aleph_\xi$ nächstfolgende Mächtigkeit nennen, da es von solchen mehrere, untereinander unvergleichbare geben könnte.

Insbesondere ist $\aleph_{\alpha+1}$ eine auf $\aleph_\alpha$ unmittelbar folgende Mächtigkeit; im Gegensatz dazu ließ uns die früher gefundene Ungleichung $2^\mathfrak{a} > \mathfrak{a}$ völlig im Ungewissen, ob zwischen diesen beiden Mächtigkeiten noch eine weitere existiert. Im einfachsten Fall ist die Frage, ob es eine Mächtigkeit zwischen $\aleph_0$ und $2^{\aleph_0}$ gibt, ja nichts anderes als das Kontinuumproblem (Kap. III, § 5).

Die Zahlenklasse $Z(\aleph_\alpha)$ der Ordnungszahlen von der Mächtigkeit $\aleph_\alpha$ hat eine kleinste Zahl ω_α, die wir die Anfangszahl dieser Zahlenklasse oder die zu $\aleph_\alpha$ gehörige Anfangszahl nennen; der Index α der Anfangszahl ω_α gibt zugleich den Typus der Menge

aller vorhergehenden Anfangszahlen. Die kleinste unendliche Ordnungszahl ω ist die Anfangszahl der Zahlenklasse $Z(\aleph_0)$, wäre also auch mit ω_0 zu bezeichnen; dann folgen die Anfangszahlen

$$\omega_1, \omega_2, \ldots, \omega_\omega, \omega_{\omega+1} \text{ usw.}$$

Es ist leicht einzusehen, daß der Limes einer Menge von Anfangszahlen wieder eine Anfangszahl, die Funktion $f(\alpha) = \omega_\alpha$ also eine Normalfunktion ist. Die Zahlenklasse $Z(\aleph_\alpha)$ besteht aus den Zahlen μ, für welche

$$\omega_\alpha \leqq \mu < \omega_{\alpha+1};$$

mit Benutzung unserer durchgängigen Bezeichnung $W(\alpha)$ für die Menge der Ordnungszahlen $< \alpha$ haben wir

$$Z(\aleph_\alpha) = W(\omega_{\alpha+1}) - W(\omega_\alpha),$$

woraus umgekehrt im Sinne der Addition geordneter Mengen folgt:

(1) $$W(\omega_{\alpha+1}) = W(\omega_\alpha) + Z(\aleph_\alpha).$$

Ferner ist

(2) $$W(\omega_\alpha) = W(\omega_0) + \sum_{\xi}^{W(\alpha)} Z(\aleph_\xi),$$

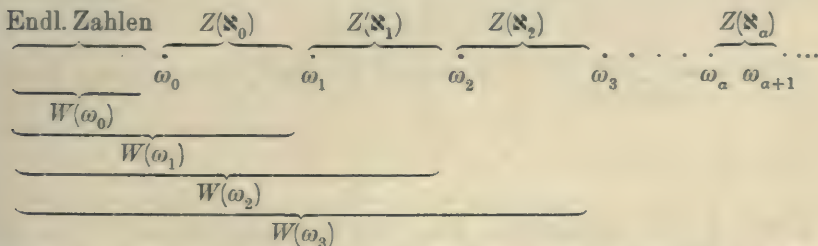
denn alle Zahlen von $W(\omega_\alpha)$ haben eine Mächtigkeit $< \aleph_\alpha$, sind also entweder endlich oder von einer der Mächtigkeiten $\aleph_\xi$ für $\xi < \alpha$. Danach ist z. B.

$$W(\omega_1) = W(\omega_0) + Z(\aleph_0),$$

d. h. $W(\omega_1)$ besteht aus den endlichen Zahlen und den Ordnungszahlen der Mächtigkeit $\aleph_0$, oder ω_1 ist der Typus, $\aleph_1$ die Mächtigkeit der Menge der Zahlen

$$0, 1, 2, \ldots, \omega, \omega + 1, \ldots, \omega 2, \ldots, \omega^2, \ldots$$

Das Schema



dürfte die Zusammenfassung der Zahlen zu Zahlenklassen und die im Zahlensystem durch die Anfangszahlen bewirkten Einschnitte einigermaßen illustrieren.[1]

---

[1] G. Cantor nennt $Z(\aleph_0)$ die zweite, $Z(\aleph_1)$ die dritte Zahlenklasse usw., indem er alle endlichen Zahlen zur ersten Zahlenklasse rechnet. Nach unserer Definition bildet jede endliche Zahl ν für sich eine Zahlenklasse $Z(\nu)$, und wir haben der Reihe nach die Klassen $Z(0)$, $Z(1)$, $Z(2)$, $\ldots$, $Z(\aleph_0)$, $Z(\aleph_1)$, $\ldots$

Nach (1) ist $Z(\aleph_\alpha)$ ein Rest (vgl. S. 103) von $W(\omega_{\alpha+1})$ und hat den Typus $-\omega_\alpha + \omega_{\alpha+1}$, ferner eine Mächtigkeit $\leqq \aleph_{\alpha+1}$ ($W(\omega_\alpha)$ hat den Typus ω_α und die Mächtigkeit $\aleph_\alpha$). Wir werden sofort zeigen, daß auch $Z(\aleph_\alpha)$ den Typus $\omega_{\alpha+1}$ und die Mächtigkeit $\aleph_{\alpha+1}$ hat.

Wir sahen gelegentlich (S. 109), daß jede Limeszahl α in der Form $\omega\beta$ darstellbar ist; für die betreffenden Mächtigkeiten folgt daraus

$$\mathfrak{a} = \aleph_0 \mathfrak{b}, \quad \aleph_0 \mathfrak{a} = \aleph_0 \aleph_0 \mathfrak{b} = \aleph_0 \mathfrak{b} = \mathfrak{a},$$

also bleibt die Mächtigkeit jeder Limeszahl bei Multiplikation mit $\aleph_0$ ungeändert.

Andererseits ist jede Anfangszahl Limeszahl; denn wäre $\omega_\alpha = \beta + 1$, so müßte β von geringerer Mächtigkeit als $\beta + 1$ sein, während doch jede unendliche Mächtigkeit $\mathfrak{b} = \mathfrak{b} + 1$ ist. Demgemäß gilt für jedes Alef

(3) $$\aleph_0 \aleph_\alpha = \aleph_\alpha$$

und um so mehr, nach dem Äquivalenzsatz,

$$n\aleph_\alpha = \aleph_\alpha \qquad (n = 1, 2, 3, \ldots)$$
$$\aleph_\alpha + \aleph_\alpha = \aleph_\alpha$$
$$\aleph_\alpha + \mathfrak{b} = \aleph_\alpha \qquad (\mathfrak{b} \leqq \aleph_\alpha)$$

Aus dieser letzten Formel ist der Schluß zu ziehen:

(4) $$\mathfrak{x} + \mathfrak{y} < \aleph_\alpha \quad \text{für } \mathfrak{x} < \aleph_\alpha,\ \mathfrak{y} < \aleph_\alpha.$$

Denn $\mathfrak{x}, \mathfrak{y}$ sind Kardinalzahlen, also jedenfalls vergleichbar; sei etwa $\mathfrak{x} \geqq \mathfrak{y}$. Ist dann $\mathfrak{x}$ ein Alef $= \aleph_\xi$ ($\xi < \alpha$), so folgt nach der letzten Formel

$$\mathfrak{x} + \mathfrak{y} = \aleph_\xi + \mathfrak{y} = \aleph_\xi < \aleph_\alpha;$$

ist $\mathfrak{x}$, also auch $\mathfrak{y}$ endlich, so ist auch $\mathfrak{x} + \mathfrak{y}$ endlich, also

$$\mathfrak{x} + \mathfrak{y} < \aleph_0 \leqq \aleph_\alpha.$$

Zerlegen wir jetzt $W(\omega_\alpha)$ irgendwie in einen Abschnitt und einen Rest, d. h. setzen wir

$$\omega_\alpha = \xi + \eta \qquad (\eta > 0)$$

und entsprechend

$$\aleph_\alpha = \mathfrak{x} + \mathfrak{y}.$$

Wegen $\xi < \omega_\alpha$ und wegen des Charakters der Anfangszahlen ist $\mathfrak{x} < \aleph_\alpha$; dann muß aber $\mathfrak{y} = \aleph_\alpha$ sein, da für $\mathfrak{y} < \aleph_\alpha$ nach (4) auch $\mathfrak{x} + \mathfrak{y} < \aleph_\alpha$ wäre. Also ist $\eta \geqq \omega_\alpha$ und zugleich $\eta \leqq \omega_\alpha$, also $\eta = \omega_\alpha$, d. h. jeder Rest von $W(\omega_\alpha)$ hat ebenfalls[1] den Typus ω_α.

[1] Diese Eigenschaft, ihren Resten gleich zu sein, kommt nicht allein den Anfangszahlen zu, sondern sämtlichen Potenzen der Basis ω und nur diesen. Den Beweis, auf Grund der S. 120 gezeigten Einschaltung einer Ordnungszahl zwischen zwei Potenzen von ω, wollen wir dem Leser überlassen.

Nach (1) hat also, wie angekündigt, $Z(\aleph_\alpha)$ den Typus $\omega_{\alpha+1}$ und die Mächtigkeit $\aleph_{\alpha+1}$; z. B. hat die Menge aller Ordnungszahlen

$$\omega,\ \omega+1,\ \ldots,\ \omega 2,\ \ldots,\ \omega^2,\ \ldots$$

von der Mächtigkeit $\aleph_0$ den Typus ω_1 und die Mächtigkeit $\aleph_1$.

Da $Z(\aleph_0)$ Teilmenge von $T(\aleph_0)$ ist, welche Typenklasse nach Kap. IV, § 7 die Mächtigkeit des Kontinuums hat, so folgt

$$2^{\aleph_0} \geqq \aleph_1;$$

die Frage, ob hier das Gleichheitszeichen gilt, ist das Kontinuumproblem.

Aus (2) folgt noch

(5) $$\omega_\alpha = \omega_0 + \sum_\xi^{W(\alpha)} \omega_{\xi+1},$$

(6) $$\aleph_\alpha = \aleph_0 + \sum_\xi^{W(\alpha)} \aleph_{\xi+1}.$$

Wir haben bereits in (3) eine Alefgleichung nebst mehreren Folgerungen kennen gelernt; die wichtigste solche Gleichung

(7) $$\aleph_\alpha^{\,2} = \aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha,$$

welche besagt, daß das Quadrat jedes Alefs ihm selbst gleich ist, müssen wir jetzt beweisen. Wir bedienen uns wieder der transfiniten Induktion. Für $\alpha = 0$ ist die Gleichung (7) richtig; wir zeigen, daß sie für α richtig ist, falls sie für alle $\xi < \alpha$ richtig ist. Der Beweis ist eine unmittelbare Übertragung jenes Diagonalverfahrens, mit dem wir eine Doppelfolge in eine einfache Folge verwandelten und damit die Gleichung $\aleph_0 \aleph_0 = \aleph_0$ bewiesen. Wir betrachten die Menge P aller geordneten Paare (ξ, η), wo $\xi < \eta < \omega_\alpha$; ordnen wir sie lexikographisch, so ist $P = \sum_\xi P_\xi$, wenn P_ξ die Menge der Paare mit festem ξ, also die Menge der Paare

$$(\xi, \xi+1)\ (\xi, \xi+2) \ldots (\xi, \eta) \ldots \qquad (\xi < \eta < \omega_\alpha)$$

bedeutet. Diese Menge ist mit $W(\omega_\alpha) - W(\xi+1)$ ähnlich, d. h. mit einem Rest von $W(\omega_\alpha)$, hat also den Typus ω_α; P hat den Typus $\sum_\xi \omega_\alpha = \omega_\alpha \omega_\alpha = \omega_\alpha^{\,2}$ und die Mächtigkeit $\aleph_\alpha^{\,2}$. Andererseits sei Q die lexikographisch geordnete Menge der umgekehrten Paare (η, ξ), also $Q = \sum_\eta Q_\eta$, wo Q_η die Menge der Paare mit festem η, d. h. der Paare

$$(\eta, 0)\ (\eta, 1) \ldots (\eta, \xi) \ldots \qquad (\xi < \eta)$$

bedeutet $(Q_0 = 0)$. Diese Menge ist mit $W(\eta)$ ähnlich, also vom Typus η, und Q vom Typus $\sum_\eta \eta$, welche Summe nach S. 107 der Limes ihrer Partialsummen

$$\beta_\eta = \sum_\xi^{W(\eta)} \xi \leqq \sum_\xi^{W(\eta)} \eta = \eta^2$$

ist. Nun ist aber $\eta^2 < \omega_\alpha$; denn vom Falle eines endlichen η abgesehen, wo diese Ungleichung trivial ist, hat η eine Mächtigkeit $\aleph_\beta < \aleph_\alpha$, η^2 die Mächtigkeit $\aleph_\beta{}^2 = \aleph_\beta < \aleph_\alpha$, weil zum Zwecke der Induktion $\aleph_\beta$ gleich seinem Quadrat angenommen war. Also ist jedes $\beta_\eta < \omega_\alpha$, der Limes dieser Zahlen β_η also $\leqq \omega_\alpha$, Q ist von einem Typus $\leqq \omega_\alpha$ und einer Mächtigkeit $\leqq \aleph_\alpha$. Aber Q ist mit P äquivalent und von der Mächtigkeit $\aleph_\alpha{}^2 \geqq \aleph_\alpha$, demnach $\aleph_\alpha{}^2 = \aleph_\alpha$ (Q ist also, beiläufig, vom Typus ω_α).

Die hiermit bewiesene Gleichung liefert nach dem assoziativen Gesetz auch

$$\aleph_\alpha{}^n = \aleph_\alpha$$

für jede natürliche Zahl n. Bezüglich weiterer Potenzen von $\aleph_\alpha$ sind wir noch ohne jede Kenntnis. Nehmen wir z. B. den Exponenten $\aleph_0$, so gibt es zweifellos unendlich viele Alefs, für die

$$\aleph_\alpha{}^{\aleph_0} > \aleph_\alpha.$$

Ein solches ist schon $\aleph_0$; ferner $\aleph_\omega$, denn nach (6) ist

$$\aleph_\omega = \aleph_0 + \aleph_1 + \aleph_2 + \aleph_3 + \cdots$$

und für die Summe

$$\mathfrak{a} = \mathfrak{a}_1 + \mathfrak{a}_2 + \mathfrak{a}_3 + \cdots$$

einer aufsteigenden Folge von Mächtigkeiten haben wir S. 59 gesehen, daß $\mathfrak{a} < \mathfrak{a}^{\aleph_0}$. Das Gleiche gilt, wie unmittelbar einleuchtet, für jedes Alef, das der Limes oder die Summe einer aufsteigenden Folge von Alefs (d. h. einer Alefmenge vom Typus ω) ist. Andererseits gibt es auch unendlich viele Mächtigkeiten, für die $\mathfrak{a}^{\aleph_0} = \mathfrak{a}$; eine solche ist z. B. die Mächtigkeit des Kontinuums $\mathfrak{a} = 2^{\aleph_0}$ und allgemein jede Potenz $\mathfrak{a} = \mathfrak{b}^{\aleph_0}$. Lassen wir also vorwegnehmend auch diese Mächtigkeiten als Alefs gelten, so gibt es auch unendlich viele Alefs, für die

$$\aleph_\alpha{}^{\aleph_0} = \aleph_\alpha.$$

Die Frage, ob $\aleph_1$ zur einen oder anderen Kategorie gehört, ob also

$$\aleph_1{}^{\aleph_0} > \aleph_1 \quad \text{oder} \quad \aleph_1{}^{\aleph_0} = \aleph_1,$$

ist wieder das Kontinuumproblem, denn wegen $\aleph_1 \leqq 2^{\aleph_0}$ ist

$$2^{\aleph_0} \leqq \aleph_1{}^{\aleph_0} \leqq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0},$$

also $\aleph_1{}^{\aleph_0}$ die Mächtigkeit des Kontinuums.

Aus der Formel (7) erfahren wir auch näheres über den Umfang der Zahlenklassen $Z(\aleph_\alpha)$ oder der Mengen $W(\omega_{\alpha+1})$. Wir können den Satz aussprechen:

I. Eine Summe von Ordnungszahlen $< \omega_{\alpha+1}$ über ein Argument $< \omega_{\alpha+1}$ ist selbst noch $< \omega_{\alpha+1}$.

Sei $\beta < \omega_{\alpha+1}$ und jeder Zahl $\eta < \beta$ eine Ordnungszahl $\mu_\eta < \omega_{\alpha+1}$ zugeordnet. Die Summe $\mu = \sum_\eta \mu_\eta$ mit dem Argument $W(\beta)$ ist dann, wie behauptet wird, auch noch $< \omega_{\alpha+1}$. Wenn $\mathfrak{m}$, $\mathfrak{m}_\eta$, $\mathfrak{b}$ die Mächtigkeiten von μ, μ_η, β sind, so ist $\mathfrak{b} < \aleph_{\alpha+1}$, also $\mathfrak{b} \leqq \aleph_\alpha$ und ebenso $\mathfrak{m}_\eta \leqq \aleph_\alpha$. Demnach ist

$$\mathfrak{m} = \sum_\eta \mathfrak{m}_\eta \leqq \sum_\eta \aleph_\alpha = \aleph_\alpha \cdot \mathfrak{b} \leqq \aleph_\alpha{}^2 = \aleph_\alpha < \aleph_{\alpha+1},$$

also wirklich $\mu < \omega_{\alpha+1}$.

Eine unmittelbare Folge davon ist:

II. Ist W eine Menge, vom Typus $\beta < \omega_{\alpha+1}$, von Ordnungszahlen $< \omega_{\alpha+1}$, so ist die nächstgrößere Zahl über W immer noch $< \omega_{\alpha+1}$.

Denn ist $W = \{\mu_0, \mu_1, \ldots, \mu_\eta, \ldots\}$, so ist die nächstgrößere Zahl über W höchstens gleich $\sum(\mu_\eta + 1) < \omega_{\alpha+1}$, da ja neben μ_η auch $\mu_\eta + 1$ kleiner als $\omega_{\alpha+1}$ ist.

Wir können auch sagen: unter den angegebenen Voraussetzungen ist $W(\omega_{\alpha+1})$ mit seiner Teilmenge W niemals konfinal.

Z. B. enthält die Menge $W(\omega_1)$ mit jeder Zahl μ auch die nächstgrößere $\mu + 1$, und mit jeder Zahlenmenge $\{\mu_0, \mu_1, \mu_2, \ldots$ vom Typus ω auch ihren Limes $\mu_\omega = \lim \mu_\nu$; dasselbe gilt auch von der Zahlenklasse $Z(\aleph_0)$. Die Menge $W(\omega_2)$ oder die Zahlenklasse $Z(\aleph_1)$ hat dieselben Eigenschaften, enthält aber außerdem mit jeder Zahlenmenge vom Typus ω_1 auch noch deren Limes; usw.

Anders als die Mengen $W(\omega_{\alpha+1})$ können sich die Mengen $W(\omega_\alpha$ verhalten. Z. B. enthält die Menge $W(\omega_\omega)$ die Zahlen $\omega_0, \omega_1, \omega_2, \ldots$, die eine Menge vom Typus ω bilden, nicht aber deren Limes ω_ω. Die Anfangszahl ω_ω ist „singulär“ (§ 6).

§ 6. Die Anfangszahlen.

Wir beweisen zunächst den wichtigen Satz:

I. Jede geordnete Menge von der Mächtigkeit $\aleph_\alpha$ ist mit einer wohlgeordneten Teilmenge vom Typus $\leqq \omega_\alpha$ konfinal.[1]

Bevor wir den Satz beweisen, einige Beispiele:

Hat die Menge ein letztes Element, so ist sie mit der aus diesem letzten Element bestehenden Teilmenge konfinal; die Menge ist mit 1 konfinal. Die Menge der rationalen Zahlen, von der

[1] Wir wollen auch sagen: die Menge A ist mit dem Typus β konfinal, wenn A mit einer Menge vom Typus β konfinal ist; der Typus α ist mit dem Typus β konfinal, wenn eine (also jede) Menge vom Typus α mit einer Menge vom Typus β konfinal ist.

Mächtigkeit $\aleph_0$, ist mit der Menge der natürlichen Zahlen, der Typus η also mit $\omega = \omega_0$ konfinal. Die Menge der reellen Zahlen, deren Mächtigkeit (unter Vorwegnahme des Wohlordnungssatzes) ein unbekanntes $\aleph_\alpha > \aleph_0$ ist, ist ebenfalls mit der Menge der natürlichen Zahlen vom Typus $\omega < \omega_\alpha$ konfinal. Die Menge $W(\omega_\omega)$, von der Mächtigkeit $\aleph_\omega$, ist mit der Teilmenge der Anfangszahlen $(\omega_0, \omega_1, \omega_2, \ldots)$ vom Typus $\omega < \omega_\omega$ konfinal.

Unsere geordnete Menge A, von der Mächtigkeit $\aleph_\alpha$, also mit $W(\omega_\alpha)$ äquivalent, läßt sich in der Form

$$A = \{a_0, a_1, \ldots, a_\xi, \ldots\} \qquad (\xi < \omega_\alpha)$$

schreiben, wobei die Ordnung der Elemente in A nicht notwendig mit der Größenordnung der Indices übereinstimmt. Es kann also, für $\xi < \eta$, $a_\xi \lesseqgtr a_\eta$ sein. Ist speziell $a_\xi < a_\eta$, und ist dies, bei einem bestimmten η, für alle $\xi < \eta$ der Fall, so nennen wir für den Augenblick a_η ein ausgezeichnetes Element; ein ausgezeichnetes Element ist also ein solches, das in A nach allen Elementen mit kleinerem Index steht. Das Element a_0 rechnen wir ebenfalls zu den ausgezeichneten Elementen. Sei nun B die Menge der ausgezeichneten Elemente. Diese Elemente haben die Ordnung ihrer Indices; B ist also wohlgeordnet und von einem Typus $\leqq \omega_\alpha$. Wir behaupten, daß A mit B konfinal ist, d. h. daß kein Element von A auf alle Elemente von B folgt. Gäbe es nämlich solche, so sei unter ihnen a_η das mit kleinstem Index. Für $\xi < \eta$ ist dann stets $a_\xi < a_\eta$, denn wäre $a_\xi > a_\eta$ für irgend ein ξ, so würde schon a_ξ auf alle Elemente von B folgen. Daraus ergibt sich aber der Widerspruch, daß a_η ein ausgezeichnetes Element ist, während es doch auf alle Elemente von B folgen, also nicht zu B gehören sollte. Hiermit ist der Satz I bewiesen.

Eine Ordnungszahl > 0, die mit einer kleineren konfinal ist, nennen wir singulär; eine solche, die mit keiner kleineren konfinal ist, regulär.

Jede endliche Zahl > 0, wie überhaupt jeder Typus einer Menge mit letztem Element, ist mit 1 konfinal, und 1 ist also die einzige endliche reguläre Zahl.

Jede unendliche reguläre Ordnungszahl ist eine Anfangszahl. Denn ist α von der Mächtigkeit $\aleph_\xi$ und $> \omega_\xi$, so ist α nach dem Satze I mit einer Zahl $\beta \leqq \omega_\xi < \alpha$ konfinal, also singulär. Unter den Zahlen von $Z(\aleph_\xi)$ kann also höchstens die Anfangszahl ω_ξ regulär sein.

Daß umgekehrt nicht jede Anfangszahl regulär ist, haben wir an ω_ω gesehen, welche Zahl mit ω konfinal, also singulär ist. Indessen gilt:

Jede Anfangszahl, deren Index keine Limeszahl ist, ist regulär. Daß $\omega_0 = \omega$ mit keiner kleineren, also endlichen Zahl konfinal sein kann, ist ja evident, da $W(\omega)$ kein letztes Element hat. Andererseits besagt der Satz § 5, II, daß $W(\omega_{\alpha+1})$ mit keiner Teilmenge W von kleinerem Typus, also $\omega_{\alpha+1}$ mit keiner kleineren Ordnungszahl konfinal ist: $\omega_{\alpha+1}$ ist regulär.

Bezüglich der Anfangszahlen mit Limesindex ist daran zu erinnern, daß ω_α eine Normalfunktion ist (S. 125); für eine Limeszahl $\alpha = \lim \xi$ ist also $\omega_\alpha = \lim \omega_\xi$ und $W(\omega_\alpha)$ mit der Menge $\{\omega_0, \omega_1, \ldots, \omega_\xi, \ldots\}$ vom Typus α konfinal. **Jede Anfangszahl ω_α mit Limesindex ist mit ihrem Index α konfinal.** Demnach ist sie singulär, falls $\omega_\alpha > \alpha$, und könnte nur regulär sein, wenn $\omega_\alpha = \alpha$, d. h. wenn α eine **kritische** Zahl für die Normalfunktion ω_α ist. Die erste dieser kritischen Zahlen ist nach § 3 der Limes der Zahlen

$$\omega = \omega_0, \quad \omega' = \omega_\omega, \quad \omega'' = \omega_{\omega'}, \quad \ldots,$$

da 0 keine kritische Zahl ($\omega_0 > 0$) ist, und diese Zahl $\varkappa = \omega_\varkappa$ ist, obwohl von einer unvorstellbar großen Mächtigkeit, doch noch singulär, da sie als Limes einer Folge mit ω konfinal ist. Wenn es also reguläre Anfangszahlen mit Limesindex gibt (und es ist bisher nicht gelungen, in dieser Annahme einen Widerspruch zu entdecken), so ist die kleinste unter ihnen von einer so exorbitanten Größe, daß sie für die üblichen Zwecke der Mengenlehre kaum jemals in Betracht kommen wird.

Unter den Ordnungszahlen, mit denen eine geordnete Menge A konfinal ist[1], ist eine die kleinste; diese muß regulär sein, denn wäre sie mit einer kleineren Zahl konfinal, so wäre A auch mit dieser konfinal (Kap. IV, § 4, I). A ist also mit einer regulären Zahl konfinal; falls ohne letztes Element, mit einer regulären Anfangszahl.

Andererseits kann A nicht zugleich mit zwei verschiedenen regulären Zahlen konfinal sein. Ist nämlich A zugleich mit der Anfangszahl ω_γ und mit $\beta < \omega_\gamma$ konfinal, so ist, wie wir sofort zeigen werden, ω_γ mit β konfinal; dann kann A aber nicht zugleich mit einer regulären Anfangszahl ω_γ und einer kleineren Zahl β, also nicht zugleich mit zwei regulären Zahlen konfinal sein. Um die ausgesprochene Behauptung zu beweisen, nehmen wir an, daß A mit einer Teilmenge B vom Typus β und mit einer Teilmenge C

[1] Vorläufig, vor dem Beweis des Wohlordnungssatzes, wäre noch die Möglichkeit zuzulassen, daß es solche Ordnungszahlen nicht gibt. Für jede Menge, die wohlgeordnet werden kann, ist die Existenz solcher Zahlen durch den Satz I gesichert.

vom Typus ω_γ konfinal ist. Die Summe $S = \mathfrak{S}(B, C)$ ist dann ebenfalls mit B und C konfinal; sie ist offenbar wohlgeordnet und von einem Typus $\geqq \omega_\gamma$. Es ist aber leicht zu sehen, daß S wie C vom Typus ω_γ ist. Sei S' der durch ein Element s bestimmte Abschnitt von S und

$$B' = \mathfrak{D}(B, S'),\ C' = \mathfrak{D}(C, S'),\ S' = \mathfrak{S}(B', C').$$

Da S mit B konfinal ist, so gibt es Elemente $b \geqq s$; ist b das erste dieser Elemente, so ist B' der durch b bestimmte Abschnitt von B. Also sind B' und C' Abschnitte von B und C, ihre Typen kleiner als β und ω_γ oder beide kleiner als ω_γ, ihre Mächtigkeiten $< \aleph_\gamma$, die Summe ihrer Mächtigkeiten nach § 5, (4) also wieder $< \aleph_\gamma$. Demnach hat auch jeder Abschnitt von S eine Mächtigkeit $< \aleph_\gamma$ und der Typus von S kann nicht $> \omega_\gamma$ sein, ist also $= \omega_\gamma$. Da S mit B konfinal ist, so ist ω_γ mit β konfinal, w. z. b. w.

Wir haben damit den Satz gewonnen:

II. Jede geordnete Menge (die wohlgeordnet werden kann) ist mit einer und nur einer regulären Ordnungszahl konfinal; falls sie kein letztes Element hat, mit einer und nur einer regulären Anfangszahl.

Dieser Satz ist für die Analysis geordneter Mengen von großer Tragweite. Speziell auf wohlgeordnete Mengen angewandt, ordnet er jeder von 0 verschiedenen Ordnungszahl α diejenige reguläre Zahl $\leqq \alpha$ zu, mit der α konfinal ist. Ist α mit 1 konfinal, so hat $W(\alpha)$ ein letztes Element oder α einen unmittelbaren Vorgänger $\alpha - 1$. Ist α mit der regulären Anfangszahl ω_ξ konfinal, so ist $W(\alpha)$ mit einer Teilmenge vom Typus ω_ξ konfinal; α ist also der Limes einer Menge vom Typus ω_ξ oder kürzer ein ω_ξ-Limes, eine ω_ξ-Zahl. Jede Zahl der Zahlenklasse $Z(\aleph_\alpha)$ ist entweder mit 1 oder mit einer regulären Anfangszahl $\leqq \omega_\alpha$ konfinal. So sind z. B. die Zahlen von $Z(\aleph_0)$ entweder Zahlen mit unmittelbarem Vorgänger oder ω-Zahlen, die Zahlen von $Z(\aleph_1)$ entweder Zahlen mit unmittelbarem Vorgänger oder ω-Zahlen oder ω_1-Zahlen usw., womit die Bemerkungen über den Umfang der Zahlenklassen am Schlusse von § 5 zu vergleichen sind. Den dortigen Sätzen I, II entsprechen hier, für beliebige reguläre Anfangszahlen, die folgenden:

III. Ist ω_α eine reguläre Anfangszahl und W eine Menge, vom Typus $\beta < \omega_\alpha$, von Ordnungszahlen $< \omega_\alpha$, so ist die nächstgrößere Zahl über W immer noch $< \omega_\alpha$.

Dies ist ja nur eine Umschreibung der Tatsache, daß $W(\omega_\alpha)$ nicht mit W, ω_α nicht mit β konfinal ist.

IV. Eine Summe von Ordnungszahlen $< \omega_\alpha$ über ein

Argument $< \omega_\alpha$ ist selbst noch $< \omega_\alpha$, wenn ω_α eine reguläre Anfangszahl ist.

Denn ist $\beta < \omega_\alpha$ und jeder Zahl $\eta < \beta$ eine Ordnungszahl $\mu_\eta < \omega_\alpha$ zugeordnet (diese brauchen nicht die Rangordnung ihrer Indices zu haben, auch nicht paarweise verschieden zu sein), so ist die Menge W der verschiedenen μ_η, der Größe nach geordnet, von einem Typus β', der höchstens dieselbe Mächtigkeit $\mathfrak{b}$ wie β hat, also $\beta' < \omega_\alpha$. Die nächstgrößere Zahl μ über W ist nach III immer noch $< \omega_\alpha$ und hat eine Mächtigkeit $\mathfrak{m} < \aleph_\alpha$. Danach ist $\sum_\eta \mu_\eta \leqq \sum_\eta \mu = \mu\beta$; die Mächtigkeit $\mathfrak{m}\mathfrak{b}$ dieses Produkts ist (von dem trivialen Fall der Endlichkeit beider Faktoren abgesehen) entweder $\leqq \mathfrak{m}^2 = \mathfrak{m} < \aleph_\alpha$ oder $\leqq \mathfrak{b}^2 = \mathfrak{b} < \aleph_\alpha$. Also ist $\mu\beta < \omega_\alpha$ und erst recht $\sum \mu_\eta < \omega_\alpha$.

Der Satz IV gilt aber auch nur für reguläre Anfangszahlen, von einem trivialen Ausnahmefall abgesehen; d. h. wenn $\alpha > 2$ und jede Summe von Zahlen $< \alpha$ über ein Argument $< \alpha$ selber $< \alpha$ ist, so ist α eine reguläre Anfangszahl.[1] Hat nämlich α einen unmittelbaren Vorgänger $\alpha - 1$, so ist eine der möglichen Summen

$$(\alpha - 1) + (\alpha - 1) > (\alpha - 1) + 1 = \alpha.$$

Ist α Limeszahl und mit $\beta < \alpha$ konfinal, so ist $W(\alpha)$ mit einer Teilmenge W vom Typus β konfinal, und die Summe aller Zahlen von W ist größer als jede Zahl von W, also $\geqq \lim W = \alpha$.

§ 7. Der Wohlordnungssatz.

Jede Menge kann wohlgeordnet werden.

G. Cantor hat diesen Satz als „Denkgesetz" ausgesprochen, aber erst E. Zermelo hat zwei strenge Beweise gegeben.

Man pflegte sich früher den Satz etwa so plausibel zu machen. Aus der unendlichen Menge A greife man willkürlich ein Element heraus, das man mit a_0 bezeichne, dann aus $A - \{a_0\}$ ein Element a_1, aus $A - \{a_0, a_1\}$ ein weiteres Element a_2 usf. Dies ist für jede endliche Zahl möglich. Wenn die Menge $\{a_0, a_1, a_2, \ldots\}$ noch nicht die ganze Menge A ist, so läßt sich aus $A - \{a_0, a_1, a_2, \ldots\}$ ein weiteres Element a_ω auswählen, wenn damit A noch nicht erschöpft ist, ein Element $a_{\omega+1}$ usw. Dies Verfahren muß einmal ein Ende nehmen, denn über der Menge W der Ordnungszahlen, denen man Elemente von A zuordnen kann, gibt es größere Zahlen, und diesen kann man also keine Elemente von A mehr zuordnen. Man kann nun

[1] Für $\alpha = 2$ ist das Argument vom Typus 1 und die beiden möglichen Summen 0 und 1 sind tatsächlich < 2. Für $\alpha = 0$ oder 1 hat der Satz überhaupt keinen Sinn.

leicht zeigen (s. u.), daß dann auch alle Elemente von A verbraucht sind, also A mit W äquivalent ist.

Wir können die meisten Bedenken, die gegen dieses Raisonnement vorgebracht worden sind, nicht teilen. Indessen bedarf die eben genannte Menge W jedenfalls einer vorsichtigeren Definition: da man von der Menge aller Ordnungszahlen nicht sprechen darf, so darf man auch nicht ohne weiteres voraussetzen, daß alle Ordnungszahlen, welche eine bestimmte Eigenschaft haben, eine Menge bilden. Ferner hat das Verfahren in der oben angegebenen Form einen unerwünschten Anschein von zeitlichem Ablauf, von dem man natürlich abstrahieren muß; in praxi wäre es ja für das menschliche Denken unmöglich, das Element a_ω erst dann auszuwählen, nachdem man schon unendlich viele Elemente $a_0, a_1, a_2, \ldots$ ausgewählt und an diese Sukzession von Wahl- oder Denkakten, deren jeder eine Minimalzeit erfordert, unendlich lange Zeit verschwendet hätte. An solche praktischen, psychologischen Bedingungen darf man sich nicht klammern: das Element a_ω ist im Sinne der transfiniten Induktion durch die Menge $W_\omega = \{a_0, a_1, a_2, \ldots\}$ der vorhergehenden Elemente bestimmt, als irgend ein Element von $A - W_\omega$, und jeden einzelnen Bestimmungsakt wie deren Reihenfolge parallel der Reihenfolge der Ordnungszahlen hat man sich gänzlich zeitlos zu denken. Zur Unterstützung dieser zeitlosen Auffassung hat E. Zermelo den glücklichen Gedanken gehabt, von vornherein aus jeder von Null verschiedenen Teilmenge A' von A eins ihrer Elemente

$$a' = f(A')$$

auszuwählen, so daß man also, sozusagen, mit dieser Auswahl nicht wartet, ob und bis die Menge A' einmal an die Reihe kommt, sondern für jede Menge, ob sie daran kommt oder nicht, das aus ihr zu wählende Element prae limine bereit hat. Das System sukzessiver Wahlakte ist damit durch ein, in der Praxis des Denkens natürlich ebenso unausführbares System simultaner Wahlakte ersetzt.

Hiernach gestaltet sich der Beweis des Wohlordnungssatzes folgendermaßen. Jeder von Null verschiedenen Teilmenge A' von A wird ein zu ihr gehöriges Element $a' = f(A')$ eindeutig zugeordnet; wir nennen dies das ausgezeichnete Element von A'. Auch die Menge A selbst hat ein ausgezeichnetes Element $f(A)$.

Ferner sei Z die Menge aller Typen[1] wohlordnungsfähiger Teilmengen von A und ζ die kleinste nicht zu Z gehörige Zahl. Ist η eine Zahl von Z und $\xi < \eta$, so ist offenbar auch ξ eine Zahl von Z;

[1] Gegen diese Menge ist nichts einzuwenden; sie ist die Summe aller Zahlenklassen $Z(\mathfrak{x})$ für $\mathfrak{x} \leqq \mathfrak{a}$, wenn $\mathfrak{a}$ die Mächtigkeit von A ist.

umgekehrt ist also jede Zahl $>\zeta$ ebenfalls nicht zu Z gehörig, während jede Zahl $<\zeta$ nach Definition zu Z gehört. Demgemäß ist $Z = W(\zeta)$. Da die Menge A selbstverständlich >0 angenommen wird, also Teilmengen vom Typus 1 hat, so ist $\zeta \geqq 2$ (für unendliches A ist leicht einzusehen, daß ζ Limeszahl ist).

Wir ordnen nun allen Elementen α von Z oder einem Teil davon Elemente a_α von A in folgender, durch transfinite Induktion bestimmter Weise zu:

Der Zahl 0 entspreche das Element $a_0 = f(A)$.

Ist $\alpha > 0$, ist jeder Zahl $\xi < \alpha$ ein Element a_ξ zugeordnet und ist die Menge W_α dieser Elemente nicht mit A identisch, so sei

$$a_\alpha = f(A - W_\alpha) = f(A_\alpha),$$

wobei $A = W_\alpha + A_\alpha$ gesetzt ist. Ist hingegen $W_\alpha = A$ oder ist einer Zahl ξ kein Element von A zugeordnet, so werde auch der Zahl α kein Element von A zugeordnet.[1]

Hiermit ist für jede Zahl α von Z entschieden, ob ihr ein Element von A entspricht oder nicht, und im ersten Falle ist dieses Element a_α eindeutig bestimmt. Aus unseren Vorschriften geht ferner hervor, daß, wenn der Zahl α ein Element a_α entspricht, auch jeder kleineren Zahl ξ ein Element a_ξ entspricht und daß $a_\alpha \neq a_\xi$ ist, da ja a_ξ zu W_α und a_α zum Komplement $A - W_\alpha$ gehört.

Sei jetzt W die Menge derjenigen Zahlen von Z, denen Elemente von A entsprechen, und α die kleinste nicht zu W gehörige Zahl. Es ist also $W \leqq Z$ und $\alpha \leqq \zeta$. Jede Zahl $>\alpha$ gehört dann ebenfalls nicht zu W, woraus analog wie bei Z hervorgeht, daß $W = W(\alpha)$. Jeder Zahl $\xi < \alpha$ entspricht ein Element a_ξ und für $\xi < \eta < \alpha$ ist $a_\xi \neq a_\eta$. Gibt man den Elementen a_ξ die Ordnung ihrer Indices, so wird $W_\alpha = \{a_0, a_1, \ldots, a_\xi, \ldots\}$ eine wohlgeordnete Teilmenge von A vom Typus α; da wir angenommen hatten, daß ζ nicht mehr Typus einer wohlgeordneten Teilmenge von A ist, so muß $\alpha < \zeta$ sein. Ferner ist dann $W_\alpha = A$, denn für $W_\alpha < A$ würde auch noch der Zahl α das Element $a_\alpha = f(A - W_\alpha)$ zuzuordnen sein. Hierdurch ist also A als wohlgeordnete Menge vom Typus α dargestellt, w. z. b. w.

Der erste Beweis von Zermelo verläuft in der Hauptsache ähnlich wie der soeben vorgetragene, vermeidet indessen die anfängliche Einführung der Menge Z und die Definition von a_α durch

---

[1] Wenn man will, kann man ein von allen Elementen von A verschiedenes Element b hinzunehmen und in dem Falle, daß der Zahl α kein Element von A entspricht, $a_\alpha = b$ setzen, womit die Funktion a_α für alle Zahlen α von Z definiert ist.

transfinite Induktion. Hier wird so geschlossen: es gibt jedenfalls wohlgeordnete Teilmengen von A

$$W_\alpha = \{a_0, a_1, \ldots, a_\xi, \ldots\} \qquad (\xi < \alpha)$$

von der Art, daß $a_\xi = f(A - W_\xi)$, insbesondere $a_0 = f(A)$; die Summe aller dieser W_α ist wieder eine solche Menge und mit A identisch. Wir begnügen uns mit dieser Charakteristik, wollen uns aber nicht versagen, den zweiten Zermeloschen Beweis hier zu reproduzieren, weil er den großen Vorzug hat, von der Theorie der Ordnungszahlen überhaupt nichts vorauszusetzen. Zum besseren Verständnis schicken wir voraus, daß er die Menge in einer ähnlichen Weise, wie sie Kap. IV, § 1 erwähnt wurde, durch Aufstellung des Systems ihrer Reste oder Endstücke ordnet.

I. Aus jeder von 0 verschiedenen Teilmenge A von M wird ein Element $a = f(A)$, das ausgezeichnete Element von A, gewählt. Die Menge

$$A' = A - \{f(A)\},$$

die nach Weglassung des ausgezeichneten Elements aus A übrig bleibt, nennen wir den Nachfolger von A.

II. Ein System $\mathfrak{S}$ von Teilmengen von M heißt eine Kette, wenn

1. M selbst zu $\mathfrak{S}$ gehört,

2. der Nachfolger A' einer zu $\mathfrak{S}$ gehörigen Menge A wieder zu $\mathfrak{S}$ gehört,

3. der Durchschnitt beliebig vieler Mengen von $\mathfrak{S}$ wieder zu $\mathfrak{S}$ gehört.

Es gibt gewiß Ketten, z. B. ist das System aller Teilmengen von M eine solche.

Der Durchschnitt beliebig vieler Ketten ist offenbar selbst eine Kette. Der Durchschnitt aller Ketten heiße die kleinste Kette $\mathfrak{K}$. Wenn also von einem Teilsystem $\mathfrak{K}' \leqq \mathfrak{K}$ sich herausstellt, daß es eine Kette ist, so muß $\mathfrak{K}' = \mathfrak{K}$ sein.

III. Ein Element A von $\mathfrak{K}$ heiße ein normales Element, wenn es alle sonstigen Elemente B von $\mathfrak{K}$ entweder als Teilmengen enthält oder als Teilmenge in ihnen enthalten ist, wenn also entweder $A > B$ oder $A < B$. Es gibt gewiß normale Elemente, z. B. M. Unterscheiden wir in bezug auf ein normales Element A die übrigen Elemente von $\mathfrak{K}$ als die Elemente $U > A$ und $V < A$.

IV. Der Nachfolger U' jedes Elements U ist $\geqq A$.

Sonst müßte nämlich $U' < A$ sein (da U' Element von $\mathfrak{K}$, also $\gtreqqless A$ ist), und aus $U > A > U'$ würde folgen, daß $U - U'$ aus mindestens zwei Elementen bestünde, während diese Menge doch nur aus dem einen Element $f(U)$ besteht.

V. Diejenigen Elemente von $\mathfrak{K}$, die $\geqq A$, und diejenigen, die $\leqq A'$, bilden zusammen eine Kette $\mathfrak{K}'$.

Die Behauptung setzt $A > 0$ voraus. Unser System $\mathfrak{K}'$ besteht aus den Elementen U, aus A und aus denjenigen Elementen W unter den V, die $\leqq A'$. Es ist zu zeigen, daß $\mathfrak{K}'$ eine Kette ist, also die Eigenschaften in II hat.

1. M ist $\geqq A$, also ein U oder A.
2. U' ist $\geqq A$ nach IV, also ein U oder A.
   A' ist ein W.
   $W' < W \leqq A'$ ist ein W.
3. Der Durchschnitt beliebig vieler Elemente von $\mathfrak{K}'$ ist, sobald unter ihnen ein W vorkommt, selbst ein W, sonst ein U oder A.

Damit ist die Behauptung bewiesen. Da nun $\mathfrak{K}' \leqq \mathfrak{K}$, so ist $\mathfrak{K}' = \mathfrak{K}$, d. h.

VI. Ist A ein normales Element > 0, so sind alle Elemente von $\mathfrak{K}$ entweder $\geqq A$ oder $\leqq A'$, also $\gtreqless A'$. Der Nachfolger eines normalen Elements ist wieder normal.

VII. Der Durchschnitt beliebig vieler normaler Elemente ist wieder normal.

Es sei $D = \mathfrak{D}(A_1, A_2, \ldots) = \mathfrak{D}_i A_i$ der Durchschnitt einer beliebigen Menge von normalen Elementen, B ein beliebiges Element von $\mathfrak{K}$. Für jedes i ist $B \gtreqless A_i$. Entweder ist, für mindestens ein i, $B \geqq A_i$, dann ist auch $B \geqq D$. Oder es ist, für jedes i, $B < A_i$, dann ist auch $B \leqq D$. Also ist D normal.

Aus VI, VII und der Tatsache, daß M normal ist, folgt, daß alle normalen Elemente von $\mathfrak{K}$ wieder eine Kette $\mathfrak{K}' \leqq \mathfrak{K}$ bilden; abermals ist also $\mathfrak{K}' = \mathfrak{K}$, mit anderen Worten:

VIII. Alle Elemente von $\mathfrak{K}$ sind normal; für zwei Elemente A, B von $\mathfrak{K}$ ist also stets eine der Relationen $A \lesseqgtr B$ erfüllt.

Hierdurch wird $\mathfrak{K}$ geordnet; wir wollen von zwei verschiedenen Elementen dasjenige als das spätere definieren, das als Teilmenge im andern enthalten ist, also $A \lesseqgtr B$ für $A \gtreqless B$. Die umfassendste Menge M ist also das erste Element von $\mathfrak{K}$.

IX. Diese Ordnung ist eine Wohlordnung. Wir haben zu beweisen, daß jedes Endstück von $\mathfrak{K}$, wenn es überhaupt Elemente hat, ein erstes Element hat. Sei $\mathfrak{K} = \mathfrak{U} + \mathfrak{V}$; U und V bezeichnen Elemente von $\mathfrak{U}$ und $\mathfrak{V}$, und es sei $U < V$, also $U > V$. Nach Annahme gibt es wirklich Elemente V und ebenso können wir die Existenz von Elementen U voraussetzen, da sonst M das erste Element von $\mathfrak{V} = \mathfrak{K}$ wäre. Es sei D der Durchschnitt aller U, also $D \geqq V$ für jedes V. Ist D selbst ein V, so ist es das erste V. Ist D ein U, also noch $D > V$, so ist D das letzte U. Dann ist der

Nachfolger D' ein V, und zwar das erste, denn gäbe es ein früheres $V > D'$, so würde aus $D > V > D'$ folgen, daß $D - D'$ aus mindestens zwei Elementen bestünde, während diese Menge ja nur das eine Element $f(D)$ hat. Also hat $\mathfrak{B}$ ein erstes Element (D oder D').

X. Es besteht eine umkehrbar eindeutige Beziehung zwischen den Elementen von M und den von 0 verschiedenen Elementen von $\mathfrak{K}$, so daß durch die Wohlordnung von $\mathfrak{K}$ auch M wohlgeordnet wird.

Jedes Element $A > 0$ von $\mathfrak{K}$ bestimmt sein ausgezeichnetes Element $a = f(A)$. Für $A \neq B$ ist $a \neq b$, denn ist etwa $A > B$, so ist $B \leqq A'$ und b gehört zu A', a hingegen nicht. Einem Element a von M kann also höchstens ein Element A von $\mathfrak{K}$ derart entsprechen, daß $a = f(A)$. Umgekehrt aber entspricht jedem a wirklich ein solches A. Ist nämlich $A = F(a)$ der Durchschnitt aller Mengen von $\mathfrak{K}$, die das Element a enthalten (zu denen z. B. M gehört), so muß $a = f(A)$ sein, da sonst $A' < A$ wäre und doch noch a enthalten würde. Die Relationen $a = f(A)$, $A = F(a)$ ordnen also jedem a umkehrbar eindeutig ein Element $A > 0$ von $\mathfrak{K}$ zu.

Definieren wir also $a < b$ für $A > B$ $\big(A = F(a),\ B = F(b)\big)$, so wird hierdurch die Menge M wohlgeordnet. Für $a < b$ ist b Element von B, also auch von A. Ist umgekehrt $a \neq b$ und b Element von A, so ist $A > B$ (denn für $A < B$ wäre $A \leqq B'$, b nicht Element von A) und $a < b$. Die Menge $A = F(a)$ stellt sich also, nach geschehener Wohlordnung von M, als Menge der Elemente $\geqq a$ heraus, d. h. als der zu a gehörige Rest, und umgekehrt ihr ausgezeichnetes Element $a = f(A)$ als das erste Element von A.

Um Mißverständnisse zu vermeiden, ist darauf aufmerksam zu machen, daß für eine beliebige, nicht zu $\mathfrak{K}$ gehörige Teilmenge A von M ihr ausgezeichnetes Element $a = f(A)$ nach der Wohlordnung nicht das erste Element von A zu sein braucht.

Mit dem Wohlordnungssatze ist nun endlich die erwünschte Einfachheit im Aufbau der Mengenlehre erreicht. Alle unendlichen Mächtigkeiten sind jetzt als Alefs und alle Mächtigkeiten als paarweise vergleichbar erkannt; $\aleph_{\alpha+1}$ ist die (nicht mehr bloß: eine) auf $\aleph_\alpha$ nächstfolgende Mächtigkeit, z. B. $\aleph_1$ die zweite unendliche Mächtigkeit. Die wirkliche Ausführung einer Wohlordnung mit Hilfe eines wirklich angegebenen Systems ausgezeichneter Elemente $a = f(A)$ entzieht sich allerdings noch vollständig unserer Fähigkeit; man überlege sich z. B. im einfachsten Fall, dem der Wohlordnung des Kontinuums (der Menge der reellen Zahlen), durch welches allgemeine Gesetz man aus jeder Menge reeller Zahlen ein Element herausgreifen solle. Infolgedessen darf man sich nicht wundern, daß wir von dem Ziel, das Kontinuum wirklich wohlzuordnen und

seiner Mächtigkeit ihren richtigen Platz unter den Alefs anzuweisen, anscheinend noch sehr weit entfernt sind (vgl. Kap. X, § 4).

Ein Beispiel, das wir dem Leser zur Übung vorschlagen: aus jeder Menge natürlicher Zahlen werde als ausgezeichnetes Element die Zahl gewählt, die die wenigsten Primfaktoren hat und unter denen mit gleicher Anzahl der Primfaktoren die kleinste ist. Die Menge der natürlichen Zahlen wird hierdurch nach dem Typus ω^2 wohlgeordnet.

Sechstes Kapitel.

Beziehungen zwischen geordneten und wohlgeordneten Mengen.

§ 1. Teilweise geordnete Mengen.

Nehmen wir an, zwischen je zwei verschiedenen Elementen a, b einer Menge A bestehe jetzt nicht mehr, wie bei geordneten Mengen, eine und nur eine von zwei Beziehungen $(a < b,\ a > b)$, sondern eine und nur eine von drei Beziehungen

$$a < b,\ a > b,\ a \| b,$$

die wir lesen wollen: a vor b, a nach b, a unvergleichbar mit b. Von den beiden ersten setzen wir dieselben Eigenschaften wie im Falle geordneter Mengen voraus, was für die dritte Beziehung notwendig ihre Symmetrie zur Folge hat, d. h.

aus $a < b,\ a > b,\ a \| b$ folgt resp. $b > a,\ b < a,\ b \| a$;

aus $a < b,\ b < c$ folgt $a < c$ (transitives Gesetz).

Eine solche Menge heiße eine *teilweise geordnete* Menge; die geordneten Mengen sind Spezialfälle der teilweise geordneten, nämlich wenn Paare unvergleichbarer Elemente nicht existieren (wozu auch der Fall zu rechnen ist, daß die Menge nur ein Element hat). Wir können auch die partielle Ordnung durch *Paarmengen* definieren, indem wir die Menge aller geordneten Paare $p = (a, b)$ von verschiedenen Elementen in drei Bestandteile P, P^*, Q spalten mit den Vorschriften:

Von zwei inversen Paaren (a, b) und (b, a) gehört *entweder* das eine zu P und das andere zu P^*, *oder* beide gehören zu Q; gehören die Paare (a, b) und (b, c) zu P, so gehört auch (a, c) zu P. Bezeichnet man dann die Zugehörigkeit eines Paares (a, b) zu

P, P^*, Q resp. durch $a < b$, $a > b$, $a \| b$, so sind die obigen Bedingungen erfüllt.

Eine teilweise geordnete Menge A hat (vollständig) geordnete Teilmengen, z. B. mindestens die aus einem Element bestehenden. Eine geordnete Teilmenge, die in keiner andern geordneten Teilmenge als echte Teilmenge enthalten ist, also nicht durch Hinzunahme anderer Elemente zu einer geordneten Teilmenge erweitert werden kann, nennen wir eine größte geordnete Teilmenge. Die Existenz solcher werden wir zu beweisen haben.

Jede nicht verschwindende Teilmenge B von A bestimmt auch hier die Menge A^B der Elemente $< B$ (die allen Elementen von B vorangehen) und die Menge A_B der Elemente $> B$. Ist $A^B = 0$ resp. $A_B = 0$, so nennen wir wieder A mit B koinitial resp. konfinal, wobei allerdings von den Sätzen I, II in Kap. IV, § 4 nur der zweite unbeschränkte Gültigkeit behält; der erste läßt sich so modifizieren:

I. Ist B eine geordnete Teilmenge von A und A mit B, B mit C koinitial, so ist auch A mit C koinitial.

Denn zu jedem Element b gibt es ein Element $c \leqq b$ (wäre B nicht geordnet, so würden wir nur schließen dürfen, daß es ein c gibt, das $= b$ oder $< b$ oder $\| b$ ist); wäre also $a < c$ für jedes c, so wäre auch $a < b$ für jedes b.

Ist B eine größte geordnete Teilmenge von A, so ist A mit B sowohl koinitial als konfinal.

Um nun die Existenz größter geordneter Teilmengen von A zu beweisen, nehmen wir den Wohlordnungssatz zu Hilfe, dessen Verfahren wir hier folgendermaßen spezialisieren (Kap. V, § 7): als ausgezeichnetes Element $a' = f(A')$ einer von Null verschiedenen Teilmenge A' von A wählen wir, wenn möglich, ein solches, das mit allen Elementen von $B' = A - A'$ vergleichbar ist ($a' \lesseqgtr b'$ in der zugrunde liegenden partiellen Ordnung von A); ist kein solches Element a' vorhanden, so wählen wir irgend ein anderes. Durch diese Wahl der ausgezeichneten Elemente wird eine Wohlordnung

$$A = \{a_0, a_1, \ldots, a_\alpha, \ldots\}$$

bestimmt, bei der

$$\begin{aligned} a_0 &= f(A), \\ a_\alpha &= f(A_\alpha) = f(A - B_\alpha), \\ B_\alpha &= \{a_0, \ldots, a_\xi, \ldots\} \qquad (\xi < \alpha). \end{aligned}$$

Wenn nun A nicht selber eine vollständig geordnete Menge ist, so muß es ein erstes Element a_α geben, das nicht mit allen früheren vergleichbar ist ($a_\alpha \| a_\xi$ für mindestens ein $\xi < \alpha$). Dagegen ist, für $\xi < \eta < \alpha$, a_η mit a_ξ vergleichbar, die Menge B_α also eine geordnete

Teilmenge von A und zwar eine größte geordnete Teilmenge, denn wäre B_α erweiterungsfähig, gäbe es also in $A_\alpha = A - B_\alpha$ Elemente, die mit allen Elementen von B_α vergleichbar sind, so wäre nach unserer Vorschrift a_α ein solches Element.

Wir haben damit für eine teilweise geordnete Menge A die Existenz größter geordneter Teilmengen B bewiesen; natürlich kann es deren verschiedene geben.[1] Es ist ferner evident, daß es zu einer gegebenen geordneten Teilmenge C von A auch mindestens eine größte geordnete Teilmenge B von A gibt, die ihrerseits C als Teilmenge enthält: um zu einer solchen zu gelangen, gebe man bezüglich der ausgezeichneten Elemente außer der obigen noch die weitere Vorschrift, daß sie, wenn möglich, der Menge C angehören sollen.

Ist B eine größte geordnete Teilmenge von A und B mit C konfinal, so ist auch A mit C konfinal (Satz I). Wir wissen aus Kap. V, § 6, II, daß B mit gewissen Ordnungszahlen, darunter mit einer regulären Zahl, konfinal ist; auch eine teilweise geordnete Menge ist also mit Ordnungszahlen, insbesondere mit regulären Zahlen konfinal, sie kann aber, im Gegensatz zu einer vollständig geordneten Menge, mit verschiedenen regulären Zahlen konfinal sein.[2]

Um eine Anwendung dieser Betrachtungen zu geben, der später noch andere folgen sollen, nehmen wir eine geordnete Menge M von mindestens zwei Elementen und deren Elementpaare $p = (a, b)$ für $a < b$. Die Menge P dieser Paare ordnen wir teilweise, indem wir folgende Vorschrift geben: es sei $p < p'$, wenn

$$a < a' < b' < b,$$

also wenn a' und b' zwischen a und b liegen; $p > p'$, wenn $p' < p$; in jedem andern Falle $p \| p'$ (falls p, p' verschieden sind, d. h. nicht gleichzeitig $a = a'$, $b = b'$ ist). Ist Q eine geordnete Teilmenge von P, vom Typus α, so sieht man unmittelbar, daß die linken Elemente a der Paare $p = (a, b)$ der Menge Q eine Teilmenge A von M vom Typus α, die rechten Elemente b eine Teilmenge B vom inversen Typus α^* bilden, während zugleich $A < B$, jedes Element von A jedem Element von B vorangeht. Ist speziell P mit Q konfinal, so

---

[1] Wenn die Vergleichbarkeit transitiv ist, d. h. zwei mit einem dritten vergleichbare Elemente auch untereinander vergleichbar sind, so zerfällt A in paarweise fremde Summanden, deren jeder eine größte geordnete Teilmenge ist.

[2] Ist z. B. $A = B + C$, B eine geordnete Menge vom Typus ω, C eine geordnete Menge vom Typus ω_1, und läßt man diesen Mengen in A ihre Ordnung, während man jedes Element b mit jedem Element c unvergleichbar annimmt, so ist A sowohl mit B als mit C, daher sowohl mit ω als auch mit ω_1 konfinal.

heißt das so viel, wie daß es zwischen allen Elementen a und allen Elementen b höchstens ein Element von M gibt; denn aus zwei solchen ließe sich ein Paar bilden, das auf alle Paare von Q folgt. Dabei kann man, wie oben gezeigt, Q als wohlgeordnet und insbesondere α als reguläre Zahl annehmen, also $\alpha = 1$ oder $\alpha = \omega_\xi$, wo ω_ξ eine reguläre Anfangszahl $\leqq \omega_\mu$ und $\aleph_\mu$ die Mächtigkeit von M (also, wegen $\aleph_\mu^2 = \aleph_\mu$, auch von P) ist. Für $\alpha = 1$ existiert also ein Paar von Elementen in M, zwischen denen ein oder kein weiteres Element liegt, d. h. es existieren benachbarte Elemente. Für $\alpha = \omega_\xi$ sind die Mengen $A < B$, von den Typen ω_ξ und ω_ξ^*, entweder benachbart und bestimmen also eine Lücke, die wir konform mit einem späteren Sprachgebrauch (§ 2) eine $\omega_\xi\omega_\xi^*$-Lücke nennen und als eine symmetrische Lücke bezeichnen, zum Unterschied von $\omega_\xi\omega_\eta^*$-Lücken mit $\omega_\xi \neq \omega_\eta$. Oder zwischen A und B liegt ein einziges Element

$$c = \lim\sup A = \lim\inf B,$$

das analog ein $\omega_\xi\omega_\xi^*$-Limes oder $\omega_\xi\omega_\xi^*$-Element oder ein symmetrischer Limes zu nennen ist. Da P mit verschiedenen regulären Zahlen konfinal sein kann, so können bei derselben Menge M verschiedene dieser Fälle gleichzeitig auftreten. Jedenfalls gilt, wenn wir die endlichen Mengen beiseite lassen, der Satz:

II. In jeder geordneten Menge von der Mächtigkeit $\aleph_\alpha$ gibt es benachbarte Elemente oder symmetrische Lücken ($\omega_\xi\omega_\xi^*$-Lücken) oder symmetrische Limites ($\omega_\xi\omega_\xi^*$-Elemente), wo ω_ξ eine reguläre Anfangszahl $\leqq \omega_\alpha$ ist.

In einer dichten Menge gibt es symmetrische Lücken oder Limites, in einer stetigen symmetrische Limites. Z. B. ist in der Menge der reellen Zahlen, in natürlicher Ordnung, jedes Element $\omega\omega^*$-Element; in der Menge der rationalen Zahlen gibt es, außer den $\omega\omega^*$-Elementen, auch $\omega\omega^*$-Lücken. In einer wohlgeordneten Menge gibt es keine Teilmengen vom Typus ω_ξ^*, also von den oben genannten drei Dingen nur benachbarte Elemente.

§ 2. Element- und Lückencharaktere.

Nach Kap. V, § 6 ist jede geordnete Menge M mit einer und nur einer regulären Zahl ϱ konfinal; ϱ ist entweder 1 (wenn die Menge ein letztes Element hat) oder eine reguläre Anfangszahl ω_ξ. Die invers geordnete Menge M^* ist ebenfalls mit einer und nur einer regulären Zahl σ ($\sigma = 1$ oder $\sigma = \omega_\eta$) konfinal, d. h. M selbst ist mit dem inversen Typus σ^* koinitial. Die beiden regulären Zahlen ϱ, σ sind durch M eindeutig bestimmt.

Diese Bemerkung kann zu speziellerer Analyse geordneter Mengen nutzbar gemacht werden. Zerlegen wir die geordnete Menge A irgendwie in zwei von 0 verschiedene Stücke

$$A = P + Q,$$

so ist P mit einer regulären Zahl ϱ konfinal, Q mit dem inversen Typus σ^* einer regulären Zahl σ koinitial; wir nennen dann das Typenpaar (ϱ, σ^*) den Charakter dieser Zerlegung $A = P + Q$. Z. B. hat die Zerlegung der natürlich geordneten Menge der reellen Zahlen

$$P = \text{Menge der Zahlen} \leqq 0, \quad Q = \text{Menge der Zahlen} > 0$$

den Charakter $(1, \omega^*)$, da P ein letztes Element hat und Q mit ω^* (z. B. mit der Menge der reellen Zahlen $1 > \frac{1}{2} > \frac{1}{3} > \ldots$) koinitial ist. Die Zerlegung

$$P = \text{Menge der Zahlen} < 0, \quad Q = \text{Menge der Zahlen} \geqq 0$$

hat den Charakter $(\omega, 1)$.

Hat die Zerlegung $P + Q$ den Charakter (ϱ, σ^*), so ist P mit einer Menge P' vom Typus ϱ konfinal, Q mit einer Menge Q' vom Typus σ^* koinitial; die Mengen $P' < Q'$ sind dann benachbart.

Sind umgekehrt $P' < Q'$ benachbarte Teilmengen, so bestimmen sie nach Kap. IV, § 4 eine Zerlegung $A = P + Q$, worin P mit P' konfinal, Q mit Q' koinitial ist. Hat diese Zerlegung den Charakter (ϱ, σ^*), so ist auch P' mit ϱ konfinal, Q' mit σ^* koinitial.

Ein Sprung ist vom Charakter $(1, 1)$, ein Schnitt vom Charakter $(\omega_\xi, 1)$ oder $(1, \omega_\eta^*)$, eine Lücke vom Charakter $(\omega_\xi, \omega_\eta^*)$.

In analoger Weise können wir den Elementen der Menge A Charaktere beilegen. Ist a ein Element von A, aber weder das erste noch das letzte, so betrachten wir die Zerlegung

$$A = P + \{a\} + Q,$$

wobei P die durch a bestimmte Anfangsstrecke, Q die Endstrecke und beide Mengen von Null verschieden sind. Ist dann P mit ϱ konfinal, Q mit σ^* koinitial (ϱ, σ reguläre Zahlen), so sagen wir, das Element a sei vom Charakter (ϱ, σ^*).

Ein Element mit unmittelbarem Vorgänger und Nachfolger ist vom Charakter $(1, 1)$. Ein Element, das nur oberer Limes ist, ist vom Charakter $(\omega_\xi, 1)$, ein Element, das nur unterer Limes ist, vom Charakter $(1, \omega_\eta^*)$. Ein Element, das beides ist, ist vom Charakter $(\omega_\xi, \omega_\eta^*)$.

Analog wie oben können wir an Stelle von P, Q die Mengen P', Q' setzen, wenn P mit P' konfinal, Q mit Q' koinitial ist.

Die Ausdrücke $\varrho\sigma^*$-Zerlegung, $\varrho\sigma^*$-Element, $\omega_\xi\omega_\eta^*$-Lücke, $\omega_\xi\omega_\eta^*$-Limes sind wohl unmittelbar verständlich.

Für $\varrho = \sigma$ sprechen wir von symmetrischen Zerlegungen, Elementen usw. Eine symmetrische Zerlegung ist entweder ein Sprung oder eine symmetrische Lücke vom Charakter $(\omega_\xi, \omega_\xi^*)$, ein symmetrisches Element hat entweder zwei Nachbarn oder ist ein symmetrischer Limes vom Charakter $(\omega_\xi, \omega_\xi^*)$.

Ist die Anfangsstrecke P von a mit ω_ξ konfinal, gleichviel wie sich Q verhält (das auch 0 sein kann[1]), so nennen wir, wie früher, a einen ω_ξ-Limes oder ein ω_ξ-Element; ist Q mit ω_η^* koinitial, einen ω_η^*-Limes oder ein ω_η^*-Element.

Sei A eine offene dichte Menge. Jedes Element hat dann einen Charakter $(\omega_\xi, \omega_\eta^*)$, und unter den Zerlegungen ziehen wir nur die Lücken in Betracht. Wir haben dann also nur Charaktere aus regulären Anfangszahlen gebildet zu berücksichtigen und schreiben dafür etwas bequemer

$$c_{\xi\eta} = (\omega_\xi, \omega_\eta^*).$$

Die Menge der Elementcharaktere nennen wir U, die Menge der Lückencharaktere (die im Fall einer stetigen Menge Null ist) V; das Mengenpaar (U, V) bezeichnen wir als Spezies und die Summe beider Mengen

$$W = \mathfrak{S}(U, V)$$

als Geschlecht der Menge A. Hiermit ist also eine Einteilung der offenen dichten Mengen nach dem Geschlecht und eine weitere Unterteilung nach der Spezies gegeben.

Für die Menge der reellen Zahlen z. B. ist, da jedes Element vom Charakter $(\omega, \omega^*) = c_{00}$ ist und Lücken nicht existieren,

$$U = \{c_{00}\}, \quad V = 0, \quad W = \{c_{00}\}.$$

Für die Menge der rationalen Zahlen ist

$$U = \{c_{00}\}, \quad V = \{c_{00}\}, \quad W = \{c_{00}\};$$

beide Mengen haben also dasselbe Geschlecht, aber verschiedene Spezies. Die Menge der irrationalen Zahlen hat dieselbe Spezies wie die der rationalen Zahlen.

Wir stellen noch einige Typen mit den zugehörigen Mengen U, V, W zusammen (η Typus der Menge der rationalen Zahlen), indem wir die Verifikation dem Leser überlassen:

$$\begin{array}{rlll}
\eta(\omega_1 + \omega_1^*): & U = \{c_{00}\}, & V = \{c_{00}, c_{11}\}, & W = \{c_{00}, c_{11}\}. \\
\eta\omega_1 + 1 + \eta\omega_1^*: & U = \{c_{00}, c_{11}\}, & V = \{c_{00}\}, & W = \{c_{00}, c_{11}\}. \\
\eta\omega_2: & U = \{c_{00}\}, & V = \{c_{00}, c_{10}\}, & W = \{c_{00}, c_{10}\}. \\
\eta\omega_2 + 1 + \eta\omega_2^*: & U = \{c_{00}, c_{22}\}, & V = \{c_{00}, c_{01}, c_{10}\}, & W = \{c_{00}, c_{01}, c_{10}, c_{22}\}.
\end{array}$$

[1] Wollten wir bei den Charakteren auch die bisher ausgeschlossenen Fälle $P = 0$, $Q = 0$ berücksichtigen, so wäre dann $\varrho = 0$, $\sigma = 0$ zu setzen; wir sehen lieber davon ab, erteilen also dem etwaigen Anfangs- oder Endelement der Menge keinen Charakter.

Die Charakterenmengen U, V, W sind nun nicht ganz beliebige Mengen, sondern müssen gewissen Bedingungen genügen. Z. B. kann W nicht etwa aus dem einen Element c_{11} bestehen. Es gäbe dann nämlich, je nachdem c_{11} Lücken- oder Elementcharakter ist, eine der beiden Zerlegungen

$$A = P + Q, \quad A = P + \{a\} + Q,$$

wo P mit ω_1 konfinal und Q mit $\omega_1{}^*$ koinitial ist. Ist P mit der Menge

$$\{a_0, a_1, \ldots, a_\omega, \ldots\}$$

vom Typus ω_1 konfinal, so ist deren Abschnitt

$$\{a_0, a_1, \ldots\}$$

vom Typus ω eine Teilmenge von A, auf deren Elemente sicher noch weitere (z. B. a_ω) folgen, und die also, je nachdem ein erstes oder kein erstes Element darauf folgt, zu einem Element oder einer Lücke vom Charakter $(\omega, \omega_\eta{}^*) = c_{0\eta}$ führen müßte. Das gleiche gilt von Q, und W kann also nicht den Charakter c_{11} enthalten, ohne auch mindestens einen Charakter $c_{0\eta}$ und einen Charakter $c_{\xi 0}$ zu enthalten.

Allgemein sieht man auf die gleiche Weise: kommt in W ein Charakter $c_{\alpha\beta}$ vor, und ist ω_ξ irgend eine reguläre Anfangszahl $< \omega_\alpha$, so muß W mindestens einen Charakter $c_{\xi\eta}$ enthalten; ist ω_η eine reguläre Anfangszahl $< \omega_\beta$, so muß W mindestens einen Charakter $c_{\xi\eta}$ enthalten.

Ist umgekehrt ω_α die kleinste reguläre Anfangszahl, für die kein Charakter $c_{\alpha\beta}$ in W vorkommt (während für jede reguläre Anfangszahl $\omega_\xi < \omega_\alpha$ ein Charakter $c_{\xi\eta}$ vorhanden ist), so kommt auch für größere reguläre Anfangszahlen kein entsprechender Charakter in W vor, und das gleiche gilt für die kleinste reguläre Anfangszahl ω_β, zu der kein Charakter $c_{\alpha\beta}$ in W vorkommt. Bildet man für alle regulären Anfangszahlen $\omega_\xi < \omega_\alpha$, $\omega_\eta < \omega_\beta$ die Charaktere $c_{\xi\eta}$ und ordnet sie in ein Tableau

$$\begin{array}{ccccc} c_{00} & c_{01} & \cdots & c_{0\eta} & \cdots \\ c_{10} & c_{11} & \cdots & c_{1\eta} & \cdots \\ \cdot & \cdot & \cdots & \cdot & \cdots \\ c_{\xi 0} & c_{\xi 1} & \cdots & c_{\xi\eta} & \cdots \\ \cdot & \cdot & \cdots & \cdot & \cdots \end{array}$$

so muß W aus jeder Zeile und Spalte dieses Tableaus mindestens ein Element enthalten.

Außerdem aber verlangt der Satz § 1, II die Existenz symmetrischer Charaktere; W muß also auch aus der „Hauptdiagonale"

$$c_{00} \quad c_{11} \quad \cdots \quad c_{\zeta\zeta} \quad \cdots$$

des Tableaus mindestens ein Element enthalten.

Es läßt sich zeigen, daß diese Bedingungen für W nicht nur notwendig, sondern auch hinreichend sind; d. h. wenn W aus jeder Zeile, aus jeder Spalte und aus der Hauptdiagonale des obigen Tableaus mindestens ein Element enthält und U, V zwei Mengen sind, deren Summe $\mathfrak{S}(U, V) = W$ und von denen natürlich U von 0 verschieden ist, so gibt es sicher offene dichte Mengen von der Spezies (U, V) und dem Geschlecht W. Wir verzichten darauf, den ziemlich komplizierten Beweis zu geben, und werden uns später mit der Konstruktion einiger besonders interessanter Mengen, augenblicklich mit einer Abzählung begnügen.

Für ein Geschlecht W mit $\mathfrak{w}$ Charakteren[1] existieren $3^{\mathfrak{w}} - 1$ Spezies. Denn setzt man

$$\mathfrak{D}(U, V) = W_0,$$
$$U = W_0 + W_1, \quad V = W_0 + W_2,$$
$$W = W_0 + W_1 + W_2,$$

so ist die Menge aller Zerlegungen $W = W_0 + W_1 + W_2$, bis auf die eine $W = 0 + 0 + W$, der Menge aller zulässigen Paare (U, V) äquivalent; die Menge aller Zerlegungen einschließlich der ausgeschiedenen hat aber die Mächtigkeit $3^{\mathfrak{w}}$.

Bei gegebenen Zahlen α, β mit den Mächtigkeiten $\mathfrak{a}, \mathfrak{b}$ hat die Menge der Spezies eine Mächtigkeit

$$s(\alpha, \beta) \geqq 3^{\mathfrak{a}\mathfrak{b}} - 1.$$

Denn das ganze obige Tableau, mit $\mathfrak{a}\mathfrak{b}$ Charakteren[2], ist jedenfalls eins der zulässigen Geschlechter. Ist also auch nur eine der Zahlen α, β unendlich, so ist die Menge der Spezies mindestens von der Mächtigkeit des Kontinuums.

Die Menge der Geschlechter hat eine Mächtigkeit

$$g(\alpha, \beta) \geqq 2^{(\mathfrak{a}-1)(\mathfrak{b}-1)}.$$

Denn die erste Zeile und Spalte des Tableaus plus einer beliebigen Teilmenge des übrig bleibenden Tableaus ist ebenfalls ein zulässiges Geschlecht. Ist von den Zahlen α, β die eine unendlich, die andere > 1, so ist also auch die Menge der Geschlechter mindestens von der Mächtigkeit des Kontinuums. Ist $\alpha = 1$ oder $\beta = 1$, so besteht das Tableau aus nur einer Zeile oder Spalte und es gibt nur ein Geschlecht.

---

[1] D. h. W hat als Charakterenmenge die Mächtigkeit $\mathfrak{w}$, die endlich oder ein Alef sein kann. Die im folgenden auftretende Bezeichnung $\mathfrak{a} - 1$ bedarf wohl keiner Erklärung, obwohl wir Subtraktion von Mächtigkeiten sonst nicht definiert haben.

[2] Es gibt ebenso viele Anfangszahlen wie reguläre Anfangszahlen $< \omega_\alpha$.

In den niedrigsten Fällen gibt es folgende Geschlechter:

$\alpha=1, \beta=1.$	$W=\{c_{00}\}$	mit	$3^1-1=2$ Spezies
$\alpha=1, \beta=2.$	$W=\{c_{00}, c_{01}\}$	„	$3^2-1=8$ „
$\alpha=2, \beta=1.$	$W=\{c_{00}, c_{10}\}$	„	$3^2-1=8$ „
$\alpha=2, \beta=2.$	$W=\{c_{00}, c_{11}\}$	„	$3^2-1=8$ „
	$W=\{c_{00}, c_{01}, c_{10}\}$	„	$3^3-1=26$ „
	$W=\{c_{00}, c_{01}, c_{11}\}$	„	$3^3-1=26$ „
	$W=\{c_{00}, c_{10}, c_{11}\}$	„	$3^3-1=26$ „
	$W=\{c_{01}, c_{10}, c_{11}\}$	„	$3^3-1=26$ „
	$W=\{c_{00}, c_{01}, c_{10}, c_{11}\}$	„	$3^4-1=80$ „

Also entsprechen diesen 4 Fällen

$$g(\alpha, \beta)=1,\ 1,\ 1,\ 6 \text{ Geschlechter}$$

$$\text{mit } s(\alpha, \beta)=2,\ 8,\ 8,\ 192 \text{ Spezies.}$$

Von Mengen, deren Element- und Lückencharaktere sich aus ω und ω_1, den beiden niedrigsten Anfangszahlen, zusammensetzen, gibt es also bereits 9 Geschlechter mit 210 Spezies, und diese Zahlen wachsen bei Zulassung höherer Charaktere äußerst rasch.

Es ist noch folgendes zu beachten. Haben die Zahlen α, β die bisherige Bedeutung, so kann die Menge A mit keiner Ordnungszahl $> \omega_\alpha$ konfinal sein, da sie sonst eine Teilmenge vom Typus ω_α mit noch weiteren darauf folgenden Elementen enthielte, welche Teilmenge einen Element- oder Lückencharakter $c_{\alpha\eta}$ bedingen würde. A ist also mit einer regulären Anfangszahl $\leqq \omega_\alpha$ konfinal und mit dem Inversen einer regulären Anfangszahl $\leqq \omega_\beta$ koinitial, wobei indessen das Gleichheitszeichen nicht auszuschließen ist; z. B. ist der Typus $\eta\omega_1$ (η Typus der Menge der rationalen Zahlen) vom Geschlecht $W=\{c_{00}\}$, aber mit ω_1 konfinal.

§ 3. Allgemeine Produkte und Potenzen.

Die Theorie der wohlgeordneten Mengen setzt uns nun auch in den Stand, die bisher auf eine endliche Zahl von Faktoren beschränkte Produktbildung geordneter Mengen zu verallgemeinern. Dies ist ebenso im Interesse der Systematik wünschenswert, wie es als Mittel zur Konstruktion von Ordnungstypen unentbehrlich ist.

Wie in Kap. II, § 2, wo es sich um ungeordnete Mengen handelte, weisen wir jedem Element i (Index) einer von Null verschiedenen geordneten Menge J, des Arguments

$$J=\{\ldots, i, \ldots, k, \ldots, l, \ldots\} \qquad (i<k<l)$$

eine von Null verschiedene geordnete Menge A_i zu und erhalten damit einen Mengenkomplex

$$(\ldots, A_i, \ldots, A_k, \ldots, A_l, \ldots).$$

Weisen wir jedem Index i ein zu A_i gehöriges Element a_i zu, so erhalten wir einen dem obigen Mengenkomplex angehörigen Elementenkomplex

$$a = (\ldots, a_i, \ldots, a_k, \ldots, a_l, \ldots).$$

Die Menge A dieser Elementenkomplexe, die allerdings zunächst ungeordnet ist, war als das Produkt

$$A = \mathfrak{P}_i^J A_i$$

der Mengen A_i definiert worden.

Wir erinnern nochmals an die lexikographische Ordnung, die wir im Falle eines endlichen J dem Produkt geben konnten. Bezeichnen wir die Elemente von J mit $1, 2, \ldots, m$, so hatten wir zwischen zwei verschiedenen Elementenkomplexen

$$a = (a_1, a_2, \ldots, a_m) \text{ und } b = (b_1, b_2, \ldots, b_m),$$

wo a_i und b_i Elemente von A_i sind, die Ordnung $a \lesseqgtr b$ festgesetzt, wenn

$$\begin{aligned} &\text{entweder} && a_1 \lesseqgtr b_1, \\ &\text{oder} && a_1 = b_1,\ a_2 \lesseqgtr b_2, \\ &\text{oder} && a_1 = b_1,\ a_2 = b_2,\ a_3 \lesseqgtr b_3 \text{ usw.} \end{aligned}$$

Mit andern Worten: ist i der erste Index, für den $a_i \neq b_i$ (während also, für $h < i$, $a_h = b_h$ ist), so soll $a \lesseqgtr b$ sein, je nachdem (in A_i) $a_i \lesseqgtr b_i$ ist; wir geben den Elementkomplexen die Ordnung ihrer ersten verschiedenen Elemente, wir ordnen sie „nach ersten Differenzen.“

Es ist vielleicht zweckmäßig, darauf aufmerksam zu machen, daß es hierbei durchaus nicht darauf ankommt, ob zwei zu verschiedenen Indices gehörige Mengen A_i, A_k gemeinsame Elemente haben oder nicht, und ob zwischen den Elementen a_i, a_k eine Ordnung besteht oder nicht; nur die Ordnung von Elementen a_i, b_i derselben Menge A_i ist auf die lexikographische Ordnung der Elementenkomplexe von Einfluß. Auch im folgenden gilt dasselbe.

Diese lexikographische Ordnung läßt sich nun zwar auf den allgemeinen Fall übertragen, aber — die Menge A der Elementenkomplexe wird dadurch im allgemeinen nur zu einer teilweise geordneten Menge (§ 1). Wenn zwei verschiedene Elementkomplexe eine erste Differenzstelle haben, d. h. wenn ein „kritischer“ Index i existiert, für den $a_i \neq b_i$, während, für $h < i$, $a_h = b_h$ ist, so können

wir $a \lesseqgtr b$ definieren, je nachdem $a_i \lesseqgtr b_i$. Es ist evident, daß aus $a < b$ zugleich $b > a$ folgt. Auch das transitive Gesetz

$$\text{aus} \quad a < b,\ b < c \quad \text{folgt}\ a < c$$

ist sofort zu beweisen. Ist nämlich

$$a_i < b_i,\ a_h = b_h \quad \text{für}\ h < i,$$
$$b_l < c_l,\ b_k = c_k \quad \text{für}\ k < l,$$

so ist

$$\begin{aligned} \text{für}\ i \leqq l: \quad & a_i < b_i,\ b_i \leqq c_i,\ \text{also}\ a_i < c_i, \\ & a_h = b_h,\ b_h = c_h,\ \text{also}\ a_h = c_h\ \text{für}\ h < i; \\ \text{für}\ i \geqq l: \quad & a_l \leqq b_l,\ b_l < c_l,\ \text{also}\ a_l < c_l, \\ & a_k = b_k,\ b_k = c_k,\ \text{also}\ a_k = c_k\ \text{für}\ k < l. \end{aligned}$$

In jedem Fall ist also $a < c$, und der kritische Index für a, c ist der frühere von den beiden kritischen Indices für a, b und b, c, falls diese verschieden sind; andernfalls mit ihnen identisch.

Wenn aber die Menge der Indices, für die $a_i \neq b_i$, kein erstes Element hat, so hätten wir a und b lexikographisch unvergleichbar ($a \| b$) zu nennen, und die Menge der Elementenkomplexe wird also im allgemeinen nur teilweise geordnet.

Wir wollen diese teilweise geordnete Menge wie früher die ungeordnete mit

$$A = \mathfrak{P}_i^J A_i$$

bezeichnen. Hier ist aber hinsichtlich der Reihenfolge der Faktoren eine zwar unbequeme, doch unvermeidliche Verabredung zu treffen. Wir sahen, daß im Falle eines endlichen

$$J = \{1, 2, \ldots, m\}$$

das lexikographisch geordnete Produkt nicht mit $A_1 A_2 \ldots A_m$, sondern mit

$$A_m \ldots A_2 A_1$$

bezeichnet werden muß, wenn wir die übliche Schreibweise von Produkten respektieren und doch die lexikographische Ordnung nicht durch die antilexikographische, nach letzten Differenzen, ersetzen wollen (diese hätte den großen Nachteil, daß wir im folgenden statt von wohlgeordneten Mengen von deren Inversionen zu sprechen hätten). Demgemäß müssen wir uns auch jetzt bei expliziter Schreibweise die Reihenfolge der Faktoren im Produkt umgekehrt wie die Reihenfolge der Indices im Argument denken; wenn wir also, wie oben, einige Indices ersichtlich machen ($i < k < l$), so ist

$$A = \mathfrak{P}_i^J A_i = \ldots A_l \ldots A_k \ldots A_i \ldots$$

zu setzen. Ist z. B. J wohlgeordnet, vom Typus ι, und speziell

$$J = W(\iota) = \{0, 1, \ldots, \eta, \ldots\} \qquad (\eta < \iota),$$

so ist

$$\overset{W(\iota)}{\underset{\eta}{\mathfrak{P}}} A_\eta = \ldots A_\eta \ldots A_1 A_0.$$

Um diese Unbequemlichkeit einigermaßen zu mildern, führen wir neben dem Argument J das inverse Argument oder den Exponenten

$$E = J^* = \{\ldots, l, \ldots, k, \ldots, i, \ldots\}$$

ein und bezeichnen das Produkt, indem wir neben dem Zeichen $\mathfrak{P}$ noch Π verwenden, mit

$$\overset{J}{\underset{i}{\mathfrak{P}}} A_i = \overset{E}{\underset{i}{\Pi}} A_i,$$

so daß die Reihenfolge der Faktoren dieselbe ist wie die der Indices im Exponenten.

Sind alle Faktoren gleich ($A_i = M$), so bezeichnen wir dementsprechend die entstehende Potenz mit

$$M^{J^*} = M^E.$$

Es hindert uns nichts, auch bei teilweise geordneten Mengen in derselben Weise wie bei geordneten Ähnlichkeit zu definieren und von ihren Typen zu sprechen, wie wir vorübergehend tun wollen. Man sieht, daß bei Ersetzung der Faktoren durch ähnliche Mengen das Produkt in ein ähnliches übergeht, sein Typus also nur von den Typen α_i der A_i abhängt; demgemäß bezeichnen wir den Typus des Produkts mit

$$\alpha = \overset{J}{\underset{i}{\mathfrak{P}}} \alpha_i = \overset{E}{\underset{i}{\Pi}} \alpha_i = \ldots \alpha_l \ldots \alpha_k \ldots \alpha_i \ldots$$

Auch die Ersetzung des Arguments durch eine ähnliche Menge in dem S. 76 präzisierten Sinne ändert den Typus des Produkts nicht. Der Typus der Potenz hängt demgemäß nur vom Typus μ der Basis M und vom Typus ι des Arguments J oder vom Typus $\varepsilon = \iota^*$ des Exponenten E ab und ist mit

$$\mu^{\iota^*} = \mu^\varepsilon$$

zu bezeichnen.

Fragen wir zunächst, wann diese Typen wirklich Ordnungstypen, Typen vollständig geordneter Mengen sind. Wenn alle A_i aus mindestens zwei Elementen bestehen, so kann man zu jeder nicht verschwindenden Teilmenge K von J zwei Elementenkomplexe a, b angeben, die sich an den Indices von K und sonst nirgends unterscheiden (d. h. K ist die Menge der i, für die $a_i \neq b_i$). Wenn also a, b stets lexikographisch vergleichbar sein sollen, so muß jedes K ein

erstes Element haben, d. h. das Argument J wohlgeordnet sein[1], was umgekehrt auch hinreicht. Im Falle wohlgeordneten Arguments läßt sich also das ganze Produkt und die ganze Potenz lexikographisch ordnen, und die genannten Typen sind Ordnungstypen, die auch die entsprechende Mächtigkeit (bei der auf die Ordnung der Faktoren nichts ankommt)

$$\mathfrak{P}_i^J \mathfrak{a}_i = \ldots \mathfrak{a}_l \ldots \mathfrak{a}_k \ldots \mathfrak{a}_i \ldots = \ldots \mathfrak{a}_i \ldots \mathfrak{a}_k \ldots \mathfrak{a}_l \ldots$$

resp. $\mathfrak{m}^\mathfrak{i} = \mathfrak{m}^\mathfrak{e}$ haben. Wir nennen sie Vollprodukte und Vollpotenzen; sie sind, wie schon durch die Bezeichnungsweise angedeutet wird, nicht mit denen in Kap. V, § 4 definierten zu verwechseln; diese letzteren haben, wie sich zeigen wird, gerade umgekehrt einen wohlgeordneten Exponenten, während wir jetzt ein wohlgeordnetes Argument voraussetzen.

Beispiele. Die Potenz ω^{ω^*} mit der Basis ω und dem Argument ω ist der Typus der Menge der Elementkomplexe

$$a = (a_0, a_1, a_2, \ldots)$$

in lexikographischer Ordnung, wo jedes a_η die Menge der natürlichen Zahlen durchläuft. Ordnen wir in derselben Weise wie Kap. III, § 5 diesem Elementenkomplex oder dieser Folge natürlicher Zahlen die reelle Zahl

$$x = \left(\frac{1}{2}\right)^{a_0} + \left(\frac{1}{2}\right)^{a_0+a_1} + \left(\frac{1}{2}\right)^{a_0+a_1+a_2} + \ldots$$

zu, wobei $0 < x \leqq 1$ und umgekehrt jeder solchen Zahl x eindeutig eine Zahlenfolge a entspricht, so erkennt man leicht, daß die lexikographische Ordnung der a die umgekehrte ist wie die natürliche Ordnung der x (für $a < b$ ist $x > y$). Da die Menge der Zahlen x in natürlicher Ordnung den Typus $\lambda + 1$ hat (λ der Typus der Menge der reellen Zahlen oder einer offenen Zahlenstrecke), so ist

$$\omega^{\omega^*} = (\lambda + 1)^* = 1 + \lambda^* = 1 + \lambda.$$

Das Produkt

$$\ldots \omega^* \omega \omega^* \omega = \mathfrak{P}_\eta^{W(\omega)} \alpha_\eta \qquad \begin{matrix}(\alpha_0 = \alpha_2 = \alpha_4 = \ldots = \omega, \\ \alpha_1 = \alpha_3 = \alpha_5 = \ldots = \omega^*)\end{matrix}$$

ist der Typus der lexikographisch geordneten Menge der Elementenkomplexe

$$a = (a_0, -a_1, a_2, -a_3, \ldots),$$

wo die a_η wiederum natürliche Zahlen in natürlicher Ordnung sind. Wir ordnen der Zahlenfolge a jetzt eindeutig den Kettenbruch

$$x = a_0 + 1 : a_1 + 1 : a_2 + 1 : a_3 + \ldots$$

[1] Wenn auch Faktoren mit nur einem Element vorkommen, so müssen die Indices der übrigen eine wohlgeordnete Teilmenge von J bilden. Der Fall, daß ein Faktor und damit das Produkt verschwindet, bleibt natürlich ausgeschlossen.

zu, der eine irrationale Zahl $x > 1$ darstellt; umgekehrt läßt sich jede irrationale Zahl $x > 1$ auf eine und nur eine Weise in einen solchen Kettenbruch entwickeln und bestimmt daher eindeutig eine Zahlenfolge a. Die lexikographische Ordnung der a ist dieselbe wie die natürliche Ordnung der x (für $a < b$ ist $x < y$). Das obige Produkt ist daher der Typus der Menge der irrationalen Zahlen > 1 oder auch aller irrationalen Zahlen.

Schon diese Beispiele zeigen, wie sich der Kreis der in Produkt- oder Potenzform darstellbaren Typen auf unserem jetzigen Standpunkt erweitert.

Gehen wir nun zu dem allgemeinen Fall über, daß das Produkt nur teilweise geordnet ist. Es liegt nahe, dann die größten geordneten Teilmengen (§ 1) in Betracht zu ziehen; indessen gibt es deren mehrere, im allgemeinen unendlich viele, und wir werden uns zwar nicht bei der lexikographischen, aber bei einer verwandten Anordnung davon überzeugen, wie wenig allgemeines man von diesen Teilmengen aussagen kann (§ 10). Es empfiehlt sich daher, spezielle geordnete Teilmengen von größerer Bestimmtheit aus der Gesamtmenge A auszusondern und als Produkte zu definieren. Daß man auf diese Weise bei gegebenem Argument und gegebenen Faktoren mehrere, von noch weiteren Zusatzbestimmungen abhängige Produkte erhält, ist bei der lexikographischen Ordnung und ihren Abarten unvermeidlich, und an deren Stelle etwas Besseres zu setzen, ist noch nicht gelungen. Auf der andern Seite kann man in der Mehrdeutigkeit der Produktbildung sogar einen Vorzug sehen, da sie den Kreis der in Produktform darstellbaren Typen erweitert.

Schicken wir voraus: für zwei Elementkomplexe a, b sei J_{ab} die Menge derjenigen Indices i, wo $a_i \neq b_i$, und $K_{ab} = J - J_{ab}$ das Komplement, also die Menge derjenigen i, wo $a_i = b_i$. Da aus $a_i = b_i$, $b_i = c_i$ auch $a_i = c_i$ folgt, so ist

$$\mathfrak{D}(K_{ab}, K_{bc}) \leqq K_{ac},$$

also umgekehrt

$$\mathfrak{S}(J_{ab}, J_{bc}) \geqq J_{ac}.$$

Zwei Elementkomplexe a, b sind nun lexikographisch vergleichbar ($a \gtreqqless b$), wenn J_{ab}, falls von 0 verschieden, ein erstes Element hat, und diese Bedingung ist reichlich erfüllt, wenn wir sogar J_{ab} als wohlgeordnet annehmen. Nennen wir in diesem Fall die beiden Elementenkomplexe a und b kongruent und schreiben

$$a \equiv b, \ b \equiv a;$$

zwei Elementkomplexe sind also kongruent, wenn sie sich nur an einer wohlgeordneten Menge von Indices unterscheiden.

Hier gilt nun das transitive Gesetz, das für die lexikographische Vergleichbarkeit schlechthin nicht gilt[1]:

$$\text{aus} \quad a \equiv b, \; b \equiv c \;\; \text{folgt} \;\; a \equiv c.$$

Denn sind J_{ab}, J_{bc} wohlgeordnete Teilmengen von J, so ist es auch ihre Summe und deren Teilmenge J_{ac}.

Sammeln wir also alle mit einem gegebenen Elementenkomplex a kongruenten zu einer Menge, die wir (A, a) nennen wollen, so sind alle Komplexe dieser Menge untereinander kongruent; für $a \equiv b$ ist $(A, a) = (A, b)$, und zwei verschiedene Mengen (A, a), (A, b) haben kein Element gemein, so daß die ganze Menge A in eine Summe solcher Bestandteile (A, a) zerfällt.

Da alle Komplexe der Menge (A, a) paarweise vergleichbar sind, so ist sie eine geordnete Teilmenge von A; überdies ist sie eine größte geordnete Teilmenge. Um dies zu zeigen, bemerken wir voraus, daß jede geordnete Menge M in zwei Komplemente P, Q zerlegt werden kann, wo P wohlgeordnet und Q ohne erstes Element ist (die Möglichkeiten $P = 0$ und $Q = 0$ sind mit zu rechnen). Zu einer solchen Zerlegung gelangt man z. B., indem man Q als Summe aller Teilmengen ($\geqq 0$) ohne erstes Element definiert; dann hat offenbar Q selber kein erstes Element und $P = M - Q$ kann keine von 0 verschiedene Teilmenge ohne erstes Element mehr haben, ist also wohlgeordnet.

Ist hiernach x ein nicht mit a kongruenter Elementenkomplex, so zerlegen wir die nicht wohlgeordnete Menge J_{ax} in der angegebenen Weise in P und Q, wobei Q von 0 verschieden ausfällt. Wir definieren dann einen Elementenkomplex b dadurch, daß für die zu P gehörigen Indices $b_i = x_i$, sonst überall $b_i = a_i$ sein soll. Dann ist, wie leicht zu sehen,

$$J_{ab} = P, \; J_{bx} = Q,$$

also $b \equiv a$ und $b \| x$, so daß x mit mindestens einem Komplex von (A, a) unvergleichbar, diese geordnete Menge also nicht erweiterungsfähig ist.

Eine solche größte geordnete Teilmenge (A, a) nennen wir nun auch ein Produkt und zwar ein Maximalprodukt; wir sagen, daß dies Produkt zum Hauptkomplex a gehört, dessen Elemente a_i wir die Hauptelemente der zugehörigen Mengen A_i nennen. Jeder Wahl der Hauptelemente oder des Hauptkomplexes entspricht ein bestimmtes Maximalprodukt, während umgekehrt jeder zu (A, a) gehörige Komplex als Hauptkomplex angesehen werden kann. Zur

---

[1] Es kann z. B. $a < b$, $b > c$ und $a \| c$ sein.

ausführlicheren Bezeichnung empfiehlt es sich, entweder nach dem Schema (A, a) der ausführlich geschriebenen Menge A den Hauptkomplex a oder auch den einzelnen Faktoren A_i ihre Hauptelemente a_i beizufügen, also zu schreiben

$$(A, a) = (\mathfrak{P}_i^J A_i, a) = \mathfrak{P}_i^J (A_i, a_i)$$
$$= (\ldots A_k \ldots A_i \ldots, a) = \ldots (A_k, a_k) \ldots (A_i, a_i) \ldots,$$

wobei noch Π_i^E für $\mathfrak{P}_i^J$ geschrieben werden kann.

Von Maximalpotenzen wollen wir in dem allgemeinen Fall $(A_i = M)$

$$(M^{J*}, a) = \mathfrak{P}_i^J (M, a_i)$$
$$= \ldots (M, a_k) \ldots (M, a_i) \ldots$$

nicht sprechen, da diese Mengen nur beschränkten Potenzcharakter haben würden, sondern nur dann, wenn auch für alle Indices das Hauptelement dasselbe Element der Basis ist $(a_i = m)$, also ein spezieller, „konstanter" Hauptkomplex

$$a = (\ldots, m, \ldots, m, \ldots)$$

vorliegt, und in diesem Fall die Potenz auch mit

$$(M, m)^{J*} = (M, m)^E$$

bezeichnen.

Um auch eine geeignete Bezeichnung für die Typen dieser Mengen zu finden, beachte man, daß das Produkt (A, a) in ein ähnliches übergeht, wenn man die Faktoren A_i durch ähnliche Mengen A_i' und dementsprechend den Elementenkomplex a durch den Elementenkomplex a' ersetzt, dessen Elemente a_i' die Bilder der Elemente a_i bei der ähnlichen Abbildung von A_i auf A_i' sind. Daraus folgt, daß der Typus von (A, a) nur von den durch die Zerlegung der Mengen

$$A_i = B_i + \{a_i\} + C_i$$

bewirkten Typenzerlegungen

$$\alpha_i = \beta_i + 1 + \gamma_i$$

(β_i der Typus der durch a_i bestimmten Anfangsstrecke B_i, γ_i der der Endstrecke C_i) oder von dem Komplex dieser sämtlichen Zerlegungen abhängt. Dies führt darauf, den Produkttypus mit

$$\mathfrak{P}_i^J (\beta_i + 1 + \gamma_i) = \Pi_i^E (\beta_i + 1 + \gamma_i)$$
$$= \ldots (\beta_l + 1 + \gamma_l) \cdots (\beta_k + 1 + \gamma_k) \cdots (\beta_i + 1 + \gamma_i) \cdots$$

und den Typus der Potenz mit

$$(\pi + 1 + \varrho)^{\iota*} = (\pi + 1 + \varrho)^{\varepsilon}$$

zu bezeichnen, wo $\mu = \pi + 1 + \varrho$ die durch das Hauptelement m bewirkte Zerlegung des Typus von $M = P + \{m\} + R$ ist. Eventuellen Zweideutigkeiten muß man hier durch Klammersetzung vorbeugen, z. B. sind die beiden Zerlegungen

$$\omega + 2 + \omega^* = \omega + 1 + (1 + \omega^*) = (\omega + 1) + 1 + \omega^*$$

zu unterscheiden.

Das wichtigste Beispiel für die Maximalprodukte liefert der Fall, daß der Exponent $E = J^*$ wohlgeordnet ist. Hier ist die Menge J_{ab} für zwei kongruente Elementenkomplexe wohlgeordnet, die inverse Menge J^*_{ab} aber auch; folglich ist sie endlich. Betrachten wir z. B. den Typus $(\pi + 1 + \varrho)^\omega$. Hier ist E vom Typus ω und kann speziell als

$$W(\omega) = \{0, 1, 2, \ldots\},$$

also

$$J = \{\ldots, 2, 1, 0\}$$

gewählt werden, und wir haben unter den lexikographisch geordneten Elementenkomplexen

$$(\ldots, x_2, x_1, x_0)$$

diejenigen zusammen zu stellen, die sich von

$$(\ldots, m, m, m)$$

nur an einer endlichen Anzahl von Stellen unterscheiden. Wieder sei $M = P + \{m\} + R$ und mit x_i, y_i, z_i bezeichnen wir beliebige Elemente von M, P, R. Unsere Potenzmenge zerfällt dann (vgl. die allgemeinen Betrachtungen in § 5) in Stücke, von denen wir Repräsentanten angeben:

$$\begin{array}{c} \vdots \\ (\ldots, m, y_3, x_2, x_1, x_0), \\ (\ldots, m, m, y_2, x_1, x_0), \\ (\ldots, m, m, m, y_1, x_0), \\ (\ldots, m, m, m, m, y_0), \\ (\ldots, m, m, m, m, m), \\ (\ldots, m, m, m, m, z_0), \\ (\ldots, m, m, m, z_1, x_0), \\ (\ldots, m, m, z_2, x_1, x_0), \\ (\ldots, m, z_3, x_2, x_1, x_0), \\ \vdots \end{array}$$

An den durch Punkte bezeichneten Stellen der Komplexe stehen Hauptelemente m; jeder Komplex geht den Komplexen der folgenden Zeilen voraus, und die Komplexe jeder Zeile (wenn x_i, y_i, z_i die Mengen M, P, R durchlaufen) sind natürlich lexikographisch zu ordnen.

Hierbei entstehen Mengen, die mit endlichen Produkten ähnlich sind, z. B. ist die Menge der Komplexe der ersten Zeile mit der Menge der Komplexe (y_3, x_2, x_1, x_0) oder mit dem Produkt $M \cdot M \cdot M \cdot P$ ähnlich und hat den Typus $\mu^3 \pi$. So erhalten wir die Formel

$$(\pi + 1 + \varrho)^{\omega} = \ldots + \mu^3\pi + \mu^2\pi + \mu\pi + \pi$$
$$+ 1 + \varrho + \mu\varrho + \mu^2\varrho + \mu^3\varrho + \ldots$$

Ist insbesondere $\pi = 0$, also das Hauptelement das erste der Basis, so ist

$$(0 + 1 + \varrho)^{\omega} = 1 + \varrho + \mu\varrho + \mu^2\varrho + \mu^3\varrho + \ldots$$

Ist, noch spezieller, $\mu = 1 + \varrho$ eine Ordnungszahl, so ist auch diese Potenz eine Ordnungszahl und zwar, für $\mu > 1$ $(\varrho > 0)$, der Limes (vgl. S. 107) der Zahlen

$$1,$$
$$1 + \varrho = \mu,$$
$$1 + \varrho + \mu\varrho = \mu + \mu\varrho = \mu^2,$$
$$1 + \varrho + \mu\varrho + \mu^2\varrho = \mu^2 + \mu^2\varrho = \mu^3 \text{ usw.},$$

also das Cantorsche μ^{ω} (Kap. V, § 4). Wir werden nachher (§ 6) allgemein beweisen, daß die Maximalpotenzen mit wohlgeordnetem Exponenten, wohlgeordneter Basis und deren erstem Element als Hauptelement nichts anderes als die damals definierten Cantorschen Potenzen sind, und das Analoge für Produkte.

Wir könnten uns mit den erhaltenen Maximalprodukten, deren Bildungsweise man als durchaus einfach und natürlich anerkennen wird, begnügen; aber es empfiehlt sich, noch weiter zu gehen und gewisse Teilmengen der (A, a) als Partialprodukte zu definieren, die dann im allgemeinen nicht mehr größte geordnete Teilmengen der Gesamtmenge A sein werden. Hierzu leitet uns gerade das letzte Beispiel, wo sich die J_{ab} als endlich herausstellten. Wir können ja jetzt einfach verlangen, daß die J_{ab} endlich sein sollen, also die Kongruenzforderung verschärfen; wenn wir diese neue Forderung durch die Schreibweise

$$a \equiv b \qquad (\omega)$$

andeuten, so ist wieder das transitive Gesetz gültig, weil aus der Endlichkeit von J_{ab} und J_{bc} sich auch die von $J_{ac} \leqq \mathfrak{S}(J_{ab}, J_{bc})$ ergibt. Diese Bemerkung führt auf den allgemeinen Ansatz, für eine Ordnungszahl μ

$$a \equiv b \qquad (\mu)$$

zu definieren[1], wenn die Menge J_{ab} der Indices, wo sich a und b unterscheiden, wohlgeordnet und von einem Typus $< \mu$ ist.

---

[1] gesprochen: a kongruent b für μ.

Indessen ist hier, wenn das transitive Gesetz gelten soll, nicht jede Wahl von μ zulässig; welche μ zulässig sind, hängt auch noch vom Typus des Arguments J ab. Allgemein und unabhängig von J sind u. a. die Anfangszahlen zulässig; denn sind J_{ab} und J_{bc} von einem Typus $< \omega_\xi$, also von einer Mächtigkeit $< \aleph_\xi$, so trifft dasselbe für ihre Summe und für J_{ac} zu. Um aber später den Beweis des assoziativen Gesetzes zu sichern, beschränken wir uns noch weiter auf reguläre Anfangszahlen ω_ξ. Sammeln wir alle mit einem gegebenen Elementkomplex a für ω_ξ kongruenten Komplexe, die also auch untereinander für ω_ξ kongruent sind, so erhalten wir eine geordnete Teilmenge $(A, a)_\xi$ von $\mathfrak{P}$, die wir ein Partialprodukt mit dem Hauptkomplex a vom Grade ξ nennen; wir hängen die Gradzahl ξ auch an die obigen ausführlichen Bezeichnungen von (A, a) rechts unten an. Für $\xi < \eta$ ist $(A, a)_\xi$ Teilmenge von $(A, a)_\eta$. Der kleinste Grad 0 entspricht der Forderung endlicher Mengen J_{ab}. Bei hinlänglich großem ξ geht das Partialprodukt in das Maximalprodukt mit demselben Hauptkomplex über; ist nämlich J von der Mächtigkeit $\aleph_\beta$, so ist jede wohlgeordnete Teilmenge von J von einem Typus $< \omega_{\beta+1}$ und $(A, a)_{\beta+1} = (A, a)$; indessen kann dies schon für einen früheren Grad eintreten. Bei wohlgeordnetem Exponenten ist schon $(A, a)_0 = (A, a)$. Endlich ist hervorzuheben, daß bei wohlgeordnetem Argument zwar nur ein einziges Maximalprodukt $(A, a) = A$ existiert, das gleich dem Vollprodukt ist, wohl aber auch hier verschiedene Partialprodukte niederer Grade existieren können.

Als Beispiel betrachten wir die Potenz

$$(0 + 1 + 1)_0^{\omega^*}$$

mit einer Basis aus zwei Elementen (die wir 0, 1 nennen), von denen das erste das Hauptelement ist, mit wohlgeordnetem Argument vom Typus ω und vom Grade 0. Man erhält diesen Typus, indem man unter den Elementkomplexen

$$a = (a_0, a_1, a_2, \ldots) \qquad (a_i = 0 \text{ oder } 1)$$

diejenigen nimmt, die nur endlich viele Einsen enthalten. Indem man dieser Zahlenfolge den dyadischen Bruch

$$x = \frac{a_0}{2} + \frac{a_1}{4} + \frac{a_2}{8} + \cdots$$

zuordnet, der eine dyadisch rationale Zahl $\geqq 0$ und < 1 darstellt, und beachtet, daß umgekehrt jedem solchen x eindeutig ein a entspricht, daß ferner die natürliche Ordnung der x mit der lexikographischen der a übereinstimmt, findet man

$$(0 + 1 + 1)_0^{\omega^*} = 1 + \eta,$$

denn die Menge der x ist abzählbar, dicht und hat ein erstes, aber kein letztes Element (Kap. IV, § 7).

Die nächste Partialpotenz ist die Vollpotenz

$$(0 + 1 + 1)_1^{\omega^*} = 2^{\omega^*},$$

denn jede Teilmenge von J ist vom Typus $\leqq \omega < \omega_1$; diese Potenz ist der Typus der lexikographisch geordneten Menge aller a. Hier entspricht jedem a als dyadischer Bruch eine reelle Zahl x $(0 \leqq x \leqq 1)$; umgekehrt entsprechen den dyadisch rationalen Zahlen x zwischen 0 und 1 (beide exklusive) zwei benachbarte Elementkomplexe, z. B. der Zahl $\frac{1}{2}$ die beiden Komplexe

$$(0, 1, 1, 1, \ldots) \text{ und } (1, 0, 0, 0, \ldots),$$

den dyadisch irrationalen Zahlen zwischen 0, 1 und den beiden Zahlen 0, 1 selbst nur ein a. Die Potenz 2^{ω^*} ist also der Typus einer Menge, die man aus der Menge der reellen Zahlen $0 \leqq x \leqq 1$ erhält, wenn man jede der Zahlen

$$\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}, \ldots$$

durch ein Paar benachbarter Elemente ersetzt denkt.

§ 4. Das assoziative Gesetz.

Es ist jetzt die naheliegende Frage zu beantworten, mit welchem Rechte wir die verschiedenen in § 3 definierten Mengen Produkte und Potenzen nennen. In erster Linie handelt es sich um den Nachweis des assoziativen Gesetzes, nämlich darum, daß man, ohne den Typus eines Produkts zu ändern, das Argument in Stücke zerlegen, die Faktoren dieser Stücke zu Zwischenprodukten zusammenfassen und aus diesen wieder ein neues Produkt zusammensetzen kann — allerdings mit bestimmten Vorschriften bezüglich der Hauptelemente der Zwischenprodukte und unter Voraussetzung gleichen Grades aller Produkte.

Zerlegen wir das Argument J in eine Summe

$$J = \sum_k^K J_k$$

mit dem zweiten Argument K und lauter von Null verschiedenen, paarweise fremden Summanden. Jeder Elementenkomplex mit den Elementen a_i bestimmt dann, wenn wir nur die Indices i betrachten, die zu J_k gehören, einen Teilkomplex b_k, und die Zuordnung dieser b_k zu den Indices k liefert einen neuen Komplex b, der sich von a eben nur durch die Zusammenfassung der Elemente in Zwischen-

komplexe unterscheidet.[1] Läßt man die a_i die Mengen A_i durchlaufen, so durchläuft a das Produkt $\mathfrak{P}_i^J A_i$, b_k das Produkt $\mathfrak{P}_i^{J_k} A_i$, b das Produkt dieser Produkte

$$\mathfrak{P}_k^K \mathfrak{P}_i^{J_k} A_i,$$

dessen Äquivalenz mit $\mathfrak{P}_i^J A_i$ wir bereits als das assoziative Gesetz für ungeordnete Produkte kennen.

Es seien jetzt a und x zwei Komplexe, b und y die ihnen entsprechenden neuen Komplexe; J' die Menge der Indices i, für die $a_i \neq x_i$. Wir haben dann

$$J' = \sum_k^K J_k', \quad J_k' = \mathfrak{D}(J', J_k),$$

und J_k' ist die Menge der zu J_k gehörigen Indices i, wo $a_i \neq x_i$, an denen sich also die Komplexe b_k und y_k unterscheiden. Ist $J_k' = 0$, so ist $b_k = y_k$ und umgekehrt. Die Menge K' der zweiten Indices k, wo $J_k' > 0$, ist also die Menge der Stellen k, wo $b_k \neq y_k$, und es ist

$$J' = \sum_k^{K'} J_k'.$$

Aus dieser Darstellung erkennen wir sofort:

Ist J' wohlgeordnet und von einem Typus $< \omega_\xi$ (ω_ξ eine reguläre Anfangszahl), so sind auch die J_k' und K' wohlgeordnet und $< \omega_\xi$. Umgekehrt: sind alle J_k' und K' wohlgeordnet und $< \omega_\xi$, so ist auch J' wohlgeordnet und $< \omega_\xi$. Diese Umkehrung folgt aus dem Satze Kap. V, § 6, IV, dessen Diskussion nun auch verständlich macht, warum wir uns auf reguläre Anfangszahlen ω_ξ beschränken mußten.

Lassen wir x jetzt also nur die Menge

$$\mathfrak{P}_i^J (A_i, a_i)_\xi = (\mathfrak{P}_i^J A_i, a)_\xi \tag{1}$$

der für ω_ξ mit a kongruenten Komplexe durchlaufen, so durchläuft y_k nur die Menge

$$B_k = \mathfrak{P}_i^{J_k} (A_i, a_i)_\xi = (\mathfrak{P}_i^{J_k} A_i, b_k)_\xi \tag{2}$$

der für ω_ξ mit b_k kongruenten Teilkomplexe, und y nur die Menge

$$\mathfrak{P}_k^K (B_k, b_k)_\xi = (\mathfrak{P}_k^K B_k, b)_\xi \tag{3}$$

[1] Vgl. die entsprechenden Bemerkungen beim Beweis des assoziativen Gesetzes für ungeordnete Produkte in Kap. II, § 3 (S. 39).

der für ω_ξ mit b kongruenten Komplexe. Die beiden Produkte (1) und (3) sind also äquivalent und überdies ähnlich, da die lexikographische Ordnung der x, wie unmittelbar aus der obigen Formel für J' zu schließen ist, dieselbe ist wie die lexikographische Ordnung der y, falls auch die y_k lexikographisch geordnet werden.

Das ist das allgemeine assoziative Gesetz, dem also als nähere Bestimmung des ursprünglich angegebenen Wortlauts hinzuzufügen ist, daß die Hauptkomplexe b_k der Zwischenprodukte B_k bei dem zweiten Produkt als Hauptelemente b_k der Faktoren B_k figurieren, und daß alle Produkte von gleichem Grade zu wählen sind.

Wir notieren einige Spezialfälle. Besteht K nur aus zwei Elementen 1, 2, so ist das zweite Produkt das Vollprodukt $B_2 B_1$, also für $J = J_1 + J_2$

$$\mathfrak{P}_i^J (A_i, a_i)_\xi \simeq \mathfrak{P}_i^{J_2} (A_i, a_i)_\xi \cdot \mathfrak{P}_i^{J_1} (A_i, a_i)_\xi .$$

Wenn J ein erstes Element hat, so kann man hiernach den letzten Faktor des Produkts abtrennen und findet auch für allgemeine Produkte das distributive Gesetz bestätigt (Kap. IV, § 2).

Setzt man alle $A_i = M$, $a_i = m$, so daß es sich um Potenzen der Basis M mit dem Hauptelement m handelt, und führt die Exponenten ein

$$F = K^*, \; E_k = J_k^*,$$

$$E = J^* = (\sum_k^K J_k)^* = \sum_k^F E_k,$$

so wird

$$(M, m)_\xi^E \simeq \prod_k^F (B_k, b_k)_\xi, \tag{4}$$

$$B_k = (M, m)_\xi^{E_k}.$$

Speziell für $J = J_1 + J_2$, $E = E_2 + E_1$

$$(M, m)_\xi^E \simeq (M, m)_\xi^{E_2} \cdot (M, m)_\xi^{E_1}, \tag{5}$$

so daß also, ohne Änderung des Typus, Potenzen multipliziert werden, indem man die Exponenten (oder die Argumente, diese aber in umgekehrter Reihenfolge) addiert.

Setzt man in (4) alle J_k oder E_k ähnlich voraus,

$$J_k \simeq H, \; E_k \simeq D, \; D = H^*,$$

so wird

$$B_k \simeq (M, m)_\xi^D = N,$$

wobei der Hauptkomplex b_k dem Hauptkomplex n dieser Potenz N entspricht, der aus lauter Hauptelementen m besteht. Demgemäß wird

$$\prod_k^F (B_k, b_k)_\xi \simeq (N, n)_\xi^F,$$

und da $E = \sum_k^F E_k \simeq DF$ wird, so folgt schließlich

(6) $$(M, m)_\xi^{DF} \simeq (N, n)_\xi^F, \quad N = (M, m)_\xi^D.$$

Eine Potenz wird also potenziert, indem man die Exponenten (oder Argumente) multipliziert; wobei aber die Wahl des Hauptelementes von N zu beachten ist.

Wir haben damit die Analoga zu den Potenzformeln in Kap. II, § 3 entwickelt; eine der Formel (10) entsprechende kommutative Formel kann selbstverständlich nicht gelten.

§ 5. Beliebige Komplexmengen.

Um mit Hilfe unserer Produkte (und Potenzen) Mengen vorgeschriebener Art zu bilden, müssen wir diese Produkte noch eingehender untersuchen, namentlich feststellen, mit welchen Mengen sie konfinal und koinitial und welches die Charaktere ihrer Zerlegungen sind (§ 2). Diese Untersuchung wird allgemeiner und zugleich einfacher, wenn wir eine ganz beliebige lexikographisch geordnete Menge A von Komplexen

$$x = (\ldots, x_i, \ldots)$$

mit dem Argument J betrachten, wo x_i die Menge A_i durchläuft, also eine beliebige geordnete Teilmenge des ganzen Produkts $\mathfrak{P}_i^J A_i$. Für zwei verschiedene Komplexe x, y von A wird also nur die lexikographische Vergleichbarkeit vorausgesetzt: d. h. die Menge der Stellen, wo $x_i \neq y_i$, hat ein erstes Element i und es ist

$$x_h = y_h \quad \text{für} \quad h < i, \quad x_i \neq y_i,$$
$$x \lesseqgtr y \quad \text{für} \quad x_i \lesseqgtr y_i.$$

Wir nennen diese erste Differenzstelle wieder den kritischen Index zwischen x und y und bezeichnen ihn mit

$$i = f(x, y) \quad \text{für} \quad x < y.$$

Wir hatten festgestellt, daß für $x < y < z$

$$f(x, z) \leqq f(x, y), \quad f(x, z) \leqq f(y, z)$$

ist und in mindestens einer dieser Formeln das Gleichheitszeichen gilt.

Damit ist zunächst die durch ein bestimmtes Element x von A bewirkte Zerlegung

(1) $$A = Y + \{x\} + Z$$

leicht darzustellen. Die zu Y gehörigen Komplexe y $(y < x)$ scheiden sich nach dem kritischen Index $i = f(y, x)$, und es sei $Y^{(i)}$ die Menge der y, für die $f(y, x) = i$, natürlich $Y^{(i)} = 0$, wenn es kein solches y gibt. Da dann aus $y \leqq y'$ immer $f(y, x) \leqq f(y', x)$, aus $f(y, x) > f(y', x)$

also $y > y'$ folgt, so ist Y die Summe der Mengen $Y^{(i)}$ in der Reihenfolge ihrer Indices. Für die z gilt Analoges, aber mit Zeichenumkehr; Z ist also die Summe der $Z^{(i)}$ in umgekehrter Reihenfolge ihrer Indices oder in der Reihenfolge, welche die Indices in dem Exponenten $J^* = E$ haben. Danach ist

$$(2) \qquad Y = \sum_i^J Y^{(i)}, \quad Z = \sum_i^{J^*} Z^{(i)}.$$

Dieselben Formeln gelten, wenn x ein nicht zu A gehöriger, aber mit allen Komplexen von A lexikographisch vergleichbarer Komplex ist, wobei nur an Stelle von (1) die Formel

$$(3) \qquad A = Y + Z$$

tritt; Y ist die Menge der zu A gehörigen Komplexe y, für die $y < x$, $Y^{(i)}$ derer, für die $f(y, x) = i$, und entsprechend sind Z, $Z^{(i)}$ erklärt.

Nach diesen Vorbemerkungen müssen wir die allgemeinste Form der Zerlegungen (3) von A in zwei Stücke feststellen, die wir beide von Null verschieden annehmen. Wir verabreden folgende Bezeichnung: ist G eine nichtverschwindende Teilmenge von J, so sei

$$x_G = (\ldots, x_g, \ldots)$$

ein Komplex, der nur den Indices g von G Elemente x_g von A_g zuordnet, x_G also ein Element des Produkts $\mathfrak{P}_g^G A_g$. Für $G < J$ ist x_G nur ein Teilkomplex, der im allgemeinen in verschiedenen Weisen zu einem vollen Komplex x ergänzt werden kann (für $G = J$ ist $x_G = x$ schon ein voller Komplex). Die Menge der Komplexe von A, die an den Stellen g die vorgeschriebenen Elemente x_g tragen (also der Komplexe u, für die $u_G = x_G$), bezeichnen wir mit $[x_G]$.

Speziell verstehen wir jetzt unter G ein Anfangsstück von J. Es ist evident, daß für zwei Komplexe $u < v$ von A auch die Anfangskomplexe u_G, v_G lexikographisch vergleichbar sind und $u_G \leqq v_G$ ist. Diese Anfangskomplexe von A, Y, Z bilden also wieder lexikographisch geordnete Komplexmengen A_G, Y_G, Z_G, wobei $y_G \leqq z_G$. Es können also die Mengen Y_G, Z_G höchstens ein Element gemein haben; haben sie eins gemein, so bezeichnen wir es mit x_G und nennen G etwa für den Augenblick ein ausgezeichnetes Anfangsstück. Dann ist natürlich x_G das letzte y_G und das erste z_G, $y_G \leqq x_G \leqq z_G$. Der Leser mache sich diese Verhältnisse an der Zerlegung eines Lexikons in zwei Bände klar, wo man nach den gemeinsamen Anfangsbuchstaben der letzten Worte des ersten Bandes

und der ersten Worte des zweiten Bandes zu fragen hat, um die Trennungsstelle beider Bände zu fixieren.

Ist F ein nichtverschwindendes Anfangsstück $\subset G$ und G ein ausgezeichnetes Anfangsstück, so ist auch F eins, und zwar ist der Anfangskomplex x_F der Anfang von x_G.

Bilden wir jetzt die Menge H aller Elemente von J, die in mindestens einem ausgezeichneten Anfangsstück vorkommen, oder die Summe aller ausgezeichneten Anfangsstücke. Wir machen zunächst die allgemeinere Annahme $H \supset 0$, setzen also die Existenz ausgezeichneter Anfangsstücke G voraus. Als Summe von Anfangsstücken ist H wieder ein Anfangsstück. Jedem Index h von H ist, da er zu einem G gehört, ein Element x_h (das für alle solche G dasselbe ist) zugeordnet, nämlich das entsprechende Element von x_G. Der Komplex

$$x_H = (\ldots, x_h, \ldots)$$

dieser Elemente ist wieder mit allen entsprechenden Anfangskomplexen y_H, z_H lexikographisch vergleichbar und

$$y_H \leqq x_H \leqq z_H.$$

Denn wenn etwa nicht $y_H = x_H$, also an irgend einer Stelle $y_h \neq x_h$ ist, und G ein den Index h enthaltendes ausgezeichnetes Anfangsstück ist, so ist $y_G \neq x_G$, also $y_G < x_G$; die erste Differenzstelle der beiden letzten Komplexe ist auch die erste Differenzstelle zwischen y_H und x_H, und zugleich $y_H < x_H$.

Wir erhalten jetzt folgende Darstellung: die Komplexe y scheiden sich zunächst nach dem kritischen Index $f(y_H, x_H)$, und es sei $Y^{(h)}$ die Menge derer, für die $f(y_H, x_H) = h$; dazu kommt noch eventuell, und zwar allen andern folgend, die Menge $\overline{Y}$ derjenigen, für die $y_H = x_H$. Analoges gilt für Z. So wird

$$Y = \sum_h^H Y^{(h)} + \overline{Y}, \quad Z = \overline{Z} + \sum_h^{H^*} Z^{(h)}. \tag{4}$$

Setzen wir noch

$$\overline{A} = \overline{Y} + \overline{Z} = [x_H], \tag{5}$$

das ist die Menge der mit x_H beginnenden Komplexe von A.

Anscheinend ist hiermit nichts gewonnen, sondern nur die Zerlegung (3) auf (5) zurückgeführt; indessen ist leicht zu sehen, daß diese letzte einen ganz einfachen Charakter hat. Wir unterscheiden, ob eine der Mengen $\overline{Y}$, $\overline{Z}$ verschwindet oder nicht.

Wenn eine der genannten Mengen verschwindet, so kann sein

$$\overline{Y} = \overline{Z} = \overline{A} = 0 \tag{6}$$

oder $\bar{A}$ von Null verschieden und

(7) $$\bar{Y} = 0, \ \bar{Z} = \bar{A}$$

oder

(8) $$\bar{Y} = \bar{A}, \ \bar{Z} = 0.$$

Hierzu bemerken wir: wenn $\bar{Y} = 0$ ((6) oder (7)), so ist

$$Y = \sum_h^H Y^{(h)} = \sum_h^{H_0} Y^{(h)},$$

wenn H_0 die Menge der h bedeutet, denen nichtverschwindende $Y^{(h)}$ entsprechen, also die Menge der wirklich vorhandenen Indices $f(y_H, x_H)$. Es ist leicht zu sehen, daß dann H mit H_0 konfinal ist und beide Mengen kein letztes Element haben. Denn für jedes h gibt es ein ausgezeichnetes Anfangsstück G, dem es angehört und dazu ein y mit $y_G = x_G$; da aber $y_H < x_H$ (nicht $= x_H$), so ist $f(y_H, x_H) > h$, es gibt also zu jedem h ein $h_0 > h$. Eine entsprechende Bemerkung ist über Z zu machen, falls $\bar{Z}$ verschwindet.

Seien nun $\bar{Y}$ und $\bar{Z}$ beide von Null verschieden, so daß $\bar{A}$ von der Zerlegung mit betroffen wird. Sind y, z zwei zugehörige Komplexe, für die also $y_H = z_H = x_H$, so ist der kritische Index $f(y, z)$ ein Element k des Endstücks $K = J - H$ $(J = H + K)$; es muß dann aber k das erste Element von K sein, da andernfalls sich noch weitere gemeinsame Anfangselemente ergäben und $J^k > H$ noch ein ausgezeichnetes Anfangsstück wäre. Dieser Fall bedingt also, daß K ein erstes Element k hat und für die Komplexe von $\bar{Y}$ und $\bar{Z}$ $f(y, z) = k$, also

$$y = (\ldots, x_h, \ldots, y_k, \ldots) \quad \text{und} \quad z = (\ldots, x_h, \ldots, z_k, \ldots)$$

mit $y_k < z_k$ ist. Wenn wir, für ein beliebiges Element a_k von A_k, die Menge der mit

$$(\ldots, x_h, \ldots, a_k) = (x_H, a_k)$$

beginnenden Komplexe von A oder $\bar{A}$ wieder durch Einschließung in eckige Klammern bezeichnen, so ist also

(9) $$\bar{Y} = \sum_{a_k}^{Y_k} [x_H, a_k], \ \bar{Z} = \sum_{a_k}^{Z_k} [x_H, a_k],$$

wobei

(10) $$A_k = Y_k + Z_k$$

eine Zerlegung des Produktfaktors A_k in zwei von 0 verschiedene Stücke ist (die wegen der Möglichkeit des Verschwindens einiger Summanden in (9) im allgemeinen noch auf verschiedene Weise so gewählt werden können, daß alle wirklich vorkommenden y_k zu Y_k und alle wirklich vorkommenden z_k zu Z_k gehören).

Hier reiht sich endlich noch der bisher ausgeschlossene Fall an, daß gar kein ausgezeichnetes Anfangsstück existiert; man sieht dann, daß schon für das ganze J der zuletzt für K festgestellte Fall eintreten, also J ein erstes Element k besitzen und $f(y, z) = k$ sein muß, womit wir die Zerlegung

$$(9') \qquad Y = \sum_{a_k}^{Y_k} [a_k], \quad Z = \sum_{a_k}^{Z_k} [a_k]$$

erhalten.

Wir geben für die damit beendigte Diskussion der Zerlegungen ein Beispiel, wo A aus Komplexen $u = (u_0, u_1, u_2, \ldots, u_\omega)$ mit dem Argument $J = W(\omega + 1)$ vom Typus $\omega + 1$ besteht; die Elemente sind Zahlen in natürlicher Ordnung. Dann bestehe Y aus den (von oben nach unten geordneten) Komplexen

$$\begin{array}{c} (0, u, u, u, \ldots, u) \\ (1, 0, u, u, \ldots, u) \\ (1, 1, 0, u, \ldots, u) \\ (1, 1, 1, 0, \ldots, u) \\ \cdot \;\; \cdot \;\; \cdot \end{array}$$

wo die mit u bezeichneten Elemente keine Rolle spielen und beliebig (auch verschieden) gewählt werden können, Z aus den Komplexen

$$\begin{array}{c} \cdot \;\; \cdot \;\; \cdot \\ (1, 1, 1, 2, \ldots, u) \\ (1, 1, 2, u, \ldots, u) \\ (1, 2, u, u, \ldots, u) \\ (2, u, u, u, \ldots, u). \end{array}$$

Hier ist, für jede natürliche Zahl n, $G = W(n)$ ein ausgezeichnetes Anfangsstück mit $x_G = (1, 1, \ldots, 1)$; weiter ist $H = W(\omega)$ und

$$x_H = (1, 1, 1, \ldots),$$

wobei zugleich der Fall (6) eintritt. Nimmt man in unsere Komplexmenge noch Komplexe auf, die mit x_H beginnen, so können diese, den Fällen (7), (8), (9) entsprechend, sämtlich zu Z oder sämtlich zu Y gerechnet oder auf beide Mengen verteilt werden.

Eine ähnliche Diskussion haben wir noch für die Zerlegungen $A = A + 0 = 0 + A$ anzustellen, zur Beantwortung der Frage, mit welchen Mengen A konfinal und koinitial ist. Wir beschränken uns auf $A = A + 0$.

Ist wieder $G > 0$ ein Anfangsstück von J und betrachten wir die Menge der Anfangskomplexe y_G für die Komplexe y von A, so kann es sein, daß diese (lexikographisch geordnete) Menge ein l e t z t e s

Element x_G hat; wir nennen dann wieder G ein ausgezeichnetes Anfangsstück. Nehmen wir wieder zunächst an, daß es solche G gebe, so gelangen wir wie vorhin zu ihrer Summe H und einem Komplex x_H derart, daß stets $y_H \leqq x_H$, ferner zu der Darstellung

(10) $$A = \sum_h^H A^{(h)} + \bar{A}, \quad \bar{A} = [x_H],$$

wo $A^{(h)}$ die Menge der Komplexe y von A ist, für die $f(y_H, x_H) = h$, und $\bar{A}$ die Menge derer, die mit x_H beginnen ($y_H = x_H$). Wir setzen wieder $J = H + K$.

Hat A ein letztes Element x, so ist J selbst ein ausgezeichnetes Anfangsstück, $H = J$, $x_H = x$, und $\bar{A}$ besteht aus dem einen Element x.

Hat A kein letztes Element, so kann $\bar{A}$ Null oder von Null verschieden sein. Ist $\bar{A} = 0$, so haben wir

(11) $$A = \sum_h^H A^{(h)} = \sum_h^{H_0} A^{(h)},$$

wo in der letzten Schreibung die nichtverschwindenden Summanden hervorgehoben sind, also H_0 die Menge der wirklich vorhandenen kritischen Indices $f(y_H, x_H)$ ist. Genau wie vorhin ergibt sich, daß H mit H_0 konfinal und jede dieser Mengen ohne letztes Element ist. Danach ist, wie wir im Interesse eines nachher auszusprechenden Satzes hervorheben, in diesem Fall A mit einer zu $H_0 \leqq J$ ähnlichen Menge konfinal.[1]

Ist $\bar{A}$ von Null verschieden (und ohne letztes Element), so ist sicher K von Null verschieden, da zwei Komplexe von $\bar{A}$ einen nicht zu H gehörigen kritischen Index haben. Wir müssen hier noch unterscheiden, ob K ein erstes Element hat oder nicht.

Hat K ein erstes Element k, so ist wieder, analog zu (9),

(12) $$\bar{A} = \sum_{a_k}^{A_k} [x_H, a_k] = \sum_{a_k}^{A_{k0}} [x_H, a_k],$$

und die Menge A_{k0} der Elemente a_k, denen nichtverschwindende Summanden entsprechen, kann kein letztes Element haben, da sonst $H + \{k\}$ ein ausgezeichnetes Anfangsstück wäre; A ist also jetzt mit einer zu $A_{k0} \leqq A_k$ ähnlichen Menge konfinal.

Hat K kein erstes Element, so sei x ein beliebiger Komplex von $\bar{A}$ und K_0 die Menge der kritischen Indices $f(x, z)$ für $z > x$.

---

[1] Eine Summe $\sum_I^J A_i$ mit nichtverschwindenden Summanden ist, wenn J kein letztes Element hat, mit einer zu J ähnlichen Menge konfinal (eine solche erhält man, wenn man aus jedem A_i ein Element auswählt); andernfalls ist sie natürlich mit dem letzten Summanden konfinal.

Dann ist (wie übrigens auch im vorigen Falle) K mit K_0 koinitial; denn andernfalls hätte K ein nichtverschwindendes Anfangsstück L, in dem kein $f(x, z)$ liegt, es würde also $z_{H+L} = x_{H+L}$ und $H + L$ ein ausgezeichnetes Anfangsstück von J sein. Demnach hat K_0 kein erstes Element, die umgekehrte Menge K_0^* kein letztes, und aus der Darstellung (1) mit

$$(13) \qquad Z = \sum_i^{J^*} Z^{(i)} = \sum_i^{K_0^*} Z^{(i)}$$

folgt jetzt, daß A mit einer zu $K_0^* \leqq J^*$ ähnlichen Menge konfinal ist.

Schließlich ist wieder in dem bisher beiseite gelassenen Fall, daß gar kein ausgezeichnetes Anfangsstück existiert, evident, daß dann für das ganze Argument J einer der zuletzt für K festgestellten Fälle eintreten muß; entweder hat also J ein erstes Element k und es ist

$$(12') \qquad A = \sum_k^{A_k} [a_k]$$

oder J hat kein erstes Element und die Formel (13) tritt in Kraft. Unsere letzte Diskussion hat also den Satz ergeben:

Eine beliebige lexikographisch geordnete Komplexmenge mit dem Argument J und den Faktoren A_i (d. h. eine geordnete Teilmenge des Produkts $\mathfrak{P}_i^J A_i$) ist stets mit einer Menge konfinal, die mit einer Teilmenge entweder des Arguments J oder des Exponenten $J^* = E$ oder eines Faktors A_i ähnlich ist.

Man kann darin ohne sonstige Änderung das Wort konfinal durch koinitial ersetzen.

Natürlich können mehrere dieser drei Fälle gleichzeitig eintreten, z. B. alle drei, wenn A ein letztes Element hat.

Wir geben wieder drei einfache Beispiele, wo A vom Typus ω und im Falle (α) mit dem Argument $W(\omega)$, im Falle (β) mit dem Exponenten $W(\omega)$, im Falle (γ) mit einem Faktor $W(\omega)$ ähnlich ist.

(α)	(β)	(γ)
(0, 0, 0, 0, ...)	(..., 0, 0, 0, 0)	(7, 2, 0, ...)
(1, 0, 0, 0, ...)	(..., 0, 0, 0, 1)	(7, 2, 1, ...)
(1, 1, 0, 0, ...)	(..., 0, 0, 1, 0)	(7, 2, 2, ...)
(1, 1, 1, 0, ...)	(..., 0, 1, 0, 0)	(7, 2, 3, ...)
(1, 1, 1, 1, ...)	(..., 1, 0, 0, 0)	(7, 2, 4, ...)
. . .	. . .	. . .

Der Komplex x_H ist bei (α) (1, 1, 1, 1, ...), bei (γ) (7, 2), bei (β) ist kein ausgezeichnetes Anfangsstück vorhanden.

§ 6. Zerlegungen von Produkten.

Die Betrachtungen des vorigen Paragraphen lassen sich nun leicht auf den Fall anwenden, daß die Menge A ein Produkt

$$A = \mathfrak{P}_i^J (A_i, a_i)_\xi$$

vom Grade ξ mit dem Hauptkomplex $a = (\ldots, a_i, \ldots)$ ist, also aus nur solchen und allen solchen Komplexen x besteht, die mit a für ω_ξ kongruent sind. Alle Mengen, die in den Formeln des § 5 auftraten, sind von einer der folgenden drei Formen:

(α) $[x_H]$ = Menge der Komplexe von A mit gegebenem Anfang $x_H = (\ldots, x_h, \ldots)$, wo H ein nichtverschwindendes Anfangsstück von J ist.

(β) $Y^{(h)}$ = Menge der Komplexe y von A, für die $y_H < x_H$ und der kritische Index $f(y_H, x_H) = h$ ist;

(γ) $Z^{(h)}$ = Menge der Komplexe z von A, für die $z_H > x_H$ und der kritische Index $f(x_H, z_H) = h$ ist.

Diese Mengen sind übrigens wohldefiniert (nur vielleicht Null), wenn man auch über x_H gar nichts voraussetzt; um aber zwecklose Allgemeinheit zu vermeiden, nehmen wir in den Fällen (β) und (γ) an, daß für jedes Element h von H mindestens ein Komplex in A vorhanden sei, der an den Stellen $\leqq h$ mit x_H übereinstimmt. Bezeichnet man die Menge dieser Stellen mit G, die der Stellen $i > h$ mit L, setzt also

(1) $$J = G + L, \quad G = J^h + \{h\}, \quad L = J_h,$$

so wird also $[x_G]$ von Null verschieden angenommen.

Alle diese Mengen resp. ihre Typen sind nun einfach anzugeben. Zur Abkürzung setzen wir für eine nichtverschwindende Teilmenge K von J

(2) $$A(K) = \mathfrak{P}_k^K (A_k, a_k)_\xi;$$

das ist also das Produkt nur über die Faktoren A_k erstreckt, oder die Menge der lexikographisch geordneten Teilkomplexe

$$u_K = (\ldots, u_k, \ldots),$$

die mit a_K für ω_ξ kongruent sind. Um nicht Ausnahmefälle immer besonders erwähnen zu müssen, definieren wir, für $K = 0$, $A(0)$ als eine Menge aus einem einzigen Element oder vom Typus 1, ähnlich wie eine Potenz mit dem Exponenten 0 gleich 1 gesetzt wird.

(α) Einen mit x_H beginnenden Komplex in A gibt es nur, wenn x_H mit a_H für ω_ξ kongruent ist. Also ist

$$[x_H] = 0 \quad \text{für } x_H \not\equiv a_H (\omega_\xi),$$

d. h. sobald die Menge der Differenzstellen zwischen x_H und a_H

nicht wohlgeordnet und $< \omega_\xi$ ist; sei es nun, daß sie wohlgeordnet und $\geqq \omega_\xi$ oder nicht einmal wohlgeordnet ist.

Ist aber $x_H \equiv a_H(\omega_\xi)$, so gibt es in A Komplexe, die mit x_H beginnen, und zwar bilden sie ein Produkt wie A selbst, nur daß die Faktoren A_h durch Mengen $\{x_h\}$ ersetzt sind, die aus je einem festen Element bestehen. Setzt man

$$J = H + K,$$

so lehrt die Anwendung des assoziativen Gesetzes, daß dies Produkt mit $\mathfrak{P}_k^K (A_k, a_k)$ ähnlich ist, also

$$[x_H] \simeq A(K) \quad \text{für} \quad x_H = a_H(\omega_\xi).$$

Die zunächst stillschweigend gemachte Voraussetzung $K > 0$ ist nach der obigen Verabredung entbehrlich, da für $K = 0$ $(H = J)$ $x_H = x$ ein voller zu A gehöriger Komplex wird, also beide Seiten der Formel sich auf Mengen mit einem Element reduzieren.

(β) Die Mengen $Y^{(h)}$ und $Z^{(h)}$ hängen von den durch die Elemente x_h bewirkten Zerlegungen

$$A_h = Y_h + \{x_h\} + Z_h$$

ab. Da wir, mit den Bezeichnungen (1), $[x_G] > 0$, also $x_G \equiv a_G(\omega_\xi)$ angenommen haben, so besteht die Menge $Y^{(h)}$ aus den Komplexen, die an den Stellen vor h mit x_H übereinstimmen, an der Stelle h ein (zu Y_h gehöriges) Element $y_h < x_h$ tragen und des weiteren beliebig verlaufen, natürlich unter der Einschränkung $y \equiv a\,(\omega_\xi)$. Die Menge dieser y bildet wieder ein Produkt wie A, worin nur A_h durch Y_h und die früheren Faktoren durch Mengen aus einem Element ersetzt sind. Nach dem assoziativen Gesetz erhält man also

$$Y^{(h)} \simeq A(L) \cdot Y_h = A(J_h) \cdot Y_h,$$

wobei der erste Faktor ein Produkt über die zu h gehörige Endstrecke J_h ist. Der zweite Faktor kann natürlich verschwinden (wenn x_h das erste Element von A_h ist) und damit auch $Y^{(h)}$ zum Verschwinden bringen.

(γ) Ebenso ist

$$Z^{(h)} \simeq A(L) \cdot Z_h = A(J_h) \cdot Z_h.$$

Für die Typen dieser Mengen sind die entsprechenden kleinen griechischen Buchstaben zu verwenden. Der Typus des ganzen Produkts war, wenn wir wie früher die durch die Hauptelemente bewirkten Zerlegungen mit

$$A_i = B_i + \{a_i\} + C_i$$

bezeichnen, gleich

$$\alpha = \mathfrak{P}_i^J (\beta_i + 1 + \gamma_i)_\xi$$

gesetzt worden; für die Teilprodukte ist

(3) $$\alpha(K) = \mathfrak{P}_k^K (\beta_k + 1 + \gamma_k)_\xi \quad \text{resp.} \quad \alpha(0) = 1.$$

Nach den Formeln in § 5 und den jetzt gewonnenen hat nunmehr jede Zerlegung

$$\alpha = \eta + \zeta$$

in zwei von Null verschiedene Summanden folgende Form. Es entspricht ihr ein Anfangsstück H von $J = H + K$ mit den Zerlegungen

$$\alpha_h = \eta_h + 1 + \zeta_h$$

und dann ist

$$\eta = \sum_h^H \alpha(J_h)\eta_h + \bar{\eta},$$

$$\zeta = \bar{\zeta} + \sum_h^{H^*} \alpha(J_h)\zeta_h,$$

wobei entweder

(a) $$\bar{\eta} = \bar{\zeta} = 0 \qquad \text{für } x_H \mp a_H(\omega_\xi),$$

oder für $x_H \equiv a_H(\omega_\xi)$ einer der folgenden drei Fälle eintritt:

(b) $$\bar{\eta} = 0, \qquad \bar{\zeta} = \alpha(K);$$

(c) $$\bar{\eta} = \alpha(K), \qquad \bar{\zeta} = 0;$$

(d) $$\bar{\eta} = \alpha(J_k)\eta_k, \quad \bar{\zeta} = \alpha(J_k)\zeta_k.$$

Der letzte Fall setzt k als erstes Element von K und eine Zerlegung

$$\alpha_k = \eta_k + \zeta_k$$

in zwei von Null verschiedene Summanden voraus. Endlich kann hierbei noch $H = 0$ und

$$\eta = \alpha(J_k)\eta_k, \quad \zeta = \alpha(J_k)\zeta_k$$

sein, mit k als erstem Element von J.

Wir hatten bemerkt, daß im Falle $\bar{\eta} = 0$

$$\eta = \sum_h^H \alpha(J_h)\eta_h = \sum_h^{H_0}{}' \alpha(J_h)\eta_h$$

(letztere Schreibweise mit nichtverschwindenden Summanden) ist, wobei H mit H_0 konfinal und ohne letztes Element ist; da hierbei η mit dem Typus von H_0 konfinal ist, so sind also, für $\bar{\eta} = 0$, η und H mit derselben regulären Anfangszahl konfinal. Für $\bar{\zeta} = 0$ sind ζ^* und H mit derselben regulären Anfangszahl konfinal.

Der Fall $\bar{\eta} = \bar{\zeta} = 0$ setzt $x_H \mp a_H(\omega_\xi)$ voraus. Da aber für jedes Anfangsintervall $G = \{\ldots, h\}$ von H noch $x_G \equiv a_G(\omega_\xi)$ ist, so muß die Menge H_{ax} der Stellen, wo $x_i \mp a_i$, wohlgeordnet und vom Typus ω_ξ sein; denn jeder ihrer Abschnitte ist wohlgeordnet und $< \omega_\xi$, sie selbst also wohlgeordnet und $\leqq \omega_\xi$, in diesem Falle also $= \omega_\xi$. Überdies ist, wie leicht einzusehen, H mit H_{ax} konfinal,

also H mit ω_ξ konfinal, demnach η und ζ^* mit ω_ξ konfinal, so daß dieser Fall stets eine Lücke vom Charakter $c_{\xi\xi}$ definiert.

Für die durch einen vollen Komplex $x = a(\omega_\xi)$ bewirkte Zerlegung schreiben wir lieber, wie schon früher,

$$\alpha = \eta + 1 + \zeta,$$

$$\eta = \sum_i^J \alpha(J_i)\eta_i, \quad \zeta = \sum_i^{J^*} \alpha(J_i)\zeta_i.$$

Insbesondere bewirkt der Hauptkomplex a die Zerlegung

$$\alpha = \beta + 1 + \gamma$$

$$= \sum_i^J \alpha(J_i)\beta_i + 1 + \sum_i^{J^*} \alpha(J_i)\gamma_i,$$

die man in gewissem Sinne als eine Rekursionsformel für Produkte ansehen kann, da das ganze Produkt $\alpha = \alpha(J)$ auf Teilprodukte über Endstrecken $\alpha(J_i)$ zurückgeführt wird; freilich ist dies cum grano salis zu verstehen, da ja J_i sozusagen nicht einfacher als J selbst ist und z. B. alle J_i mit J ähnlich sein können.

Es gibt aber einen wichtigen Fall, in dem die Rekursionsformel wirklich als solche wirkt, nämlich den Fall eines wohlgeordneten Exponenten, der seinerseits wieder den Fall der Cantorschen Produkte (Kap. V, § 4) als Spezialfall einschließt. Machen wir folgende Annahmen:

(A) Der Exponent $E = J^*$ ist wohlgeordnet.

Wir können dann den Grad ξ weglassen, da die Produkte aller Grade mit dem Maximalprodukt übereinstimmen. Führen wir statt der Argumente die Exponenten ein, wobei der Endstrecke J_i als inverse Menge die Anfangsstrecke oder der Abschnitt E^i entspricht, und das Produktzeichen $\mathfrak{P}$ durch Π zu ersetzen ist, so hat man

$$\beta = \sum_i^{E^*} \prod_k^{E^i} (\beta_k + 1 + \gamma_k) \cdot \beta_i,$$

$$\gamma = \sum_i^{E} \prod_k^{E^i} (\beta_k + 1 + \gamma_k) \cdot \gamma_i.$$

Sei E vom Typus h und speziell[1]

$$E = W(h) = \{0, 1, \ldots, i, \ldots\}, \qquad (i < h)$$

$$E^i = W(i) = \{0, 1, \ldots, k, \ldots\}. \qquad (k < i)$$

Indem wir die Zerlegungen $\beta_i + 1 + \gamma_i$ als gegeben ansehen, wollen

[1] Um nicht neue Buchstaben einzuführen, verstehen wir also jetzt unter $h > i > k$ Ordnungszahlen.

wir die auftretenden Produkte in ihrer Abhängigkeit vom Exponenten kurz mit

$$f(i) = \prod_k^{W(i)} (\beta_k + 1 + \gamma_k) = (\beta_0 + 1 + \gamma_0)(\beta_1 + 1 + \gamma_1) \cdots (\beta_k + 1 + \gamma_k) \cdots,$$

das ganze Produkt analog mit $f(h)$ bezeichnen. Wir haben also

$$f(h) = \beta + 1 + \gamma,$$

$$\beta = \sum_i^{W(h)*} f(i)\beta_i, \quad \gamma = \sum_i^{W(h)} f(i)\gamma_i$$

oder

$$\beta = \ldots + f(i)\beta_i + \ldots + f(1)\beta_1 + f(0)\beta_0,$$

$$\gamma = f(0)\gamma_0 + f(1)\gamma_1 + \ldots + f(i)\gamma_i + \ldots$$

(wobei $f(0) = 1$).

(B) Alle Hauptelemente sind die ersten Elemente ihrer Mengen. Es sind also alle $\beta_i = 0$ und es handelt sich um die Produkte

$$f(i) = \prod_k^{W(i)} (0 + 1 + \gamma_k) = (0 + 1 + \gamma_0)(0 + 1 + \gamma_1) \ldots (0 + 1 + \gamma_k) \ldots$$

Dann ist auch $\beta = 0$ und

$$f(h) = 1 + \sum_i^{W(h)} f(i)\gamma_i. \tag{4}$$

(C) Alle Faktoren sind wohlgeordnet.

Die $\alpha_i = 1 + \gamma_i$ sind also Ordnungszahlen. Durch Induktion folgt aus (4), daß dann auch die Produkte $f(h)$ Ordnungszahlen sind. Ferner ist nach den Betrachtungen S. 107 $f(h)$ die erste Zahl, die $\geqq$ sämtlichen Zahlen

$$1 + \sum_k^{W(i+1)} f(k)\gamma_k = 1 + \sum_k^{W(i)} f(k)\gamma_k + f(i)\gamma_i$$

$$= f(i) + f(i)\gamma_i = f(i)\alpha_i$$

ist, und dies ist genau die Definition der Cantorschen Produkte

$$\prod_i^{W(h)} \alpha_i = \alpha_0\,\alpha_1 \ldots \alpha_i \ldots$$

Also: die Cantorschen Produkte sind Maximalprodukte mit wohlgeordnetem Exponenten, wohlgeordneten Faktoren und deren ersten Elementen als Hauptelementen.

Speziell: die Cantorschen Potenzen sind Maximalpotenzen mit wohlgeordnetem Exponenten, wohlgeordneter Basis und deren erstem Element als Hauptelement.

§ 7. Potenzen mit wohlgeordnetem Argument.

Der Fall wohlgeordneten Arguments ist das Gegenstück zum Fall eines wohlgeordneten Exponenten. Während dieser nur Produkte

vom Grade 0 zuläßt, die gleichzeitig Maximalprodukte sind, bietet jener wie der allgemeine Fall beliebigen Arguments in den Produkten verschiedener Grade (die schließlich bei hinlänglich hohem Grade Vollprodukte werden) eine reichere Auswahl konstruierbarer Typen; dementsprechend ist allerdings auch die Diskussion in viele Einzelfälle verzweigt. Hier eröffnet sich auch die Möglichkeit, dichte Mengen aus undichten, ja zerstreuten Faktoren zu bilden, z. B. aus endlichen oder aus solchen, die aus wohlgeordneten Mengen und deren Umkehrungen zusammengesetzt sind; denn dazu müssen jedenfalls Argumente ohne letztes Element verwendet werden.[1] Wir können sogar gewisse Homogenitätseigenschaften erzielen, wie wir sie an der Menge der reellen oder rationalen Zahlen beobachten und von dort auf den Raum der Geometrie übertragen, Eigenschaften, die aussagen, daß diese Gebilde gewisse Transformationen gestatten. Für allgemeine geordnete Mengen, deren Elemente nicht wie Zahlen solchen Verknüpfungen wie Addition usw. unterworfen werden können, müssen wir erst einmal sagen, was wir unter homogen verstehen wollen. Wir nennen eine von Null verschiedene Menge homogen, wenn sie mit allen ihren Strecken ähnlich ist. Offenbar genügt dazu die Bedingung, daß sie mit allen ihren Anfangs- und Endstrecken ähnlich ist; denn dann ist sie auch mit allen Endstrecken von Anfangsstrecken (oder umgekehrt), d. h. mit allen Mittelstrecken ähnlich. Die Typen η und λ (der Menge der rationalen und reellen Zahlen) sind die einfachsten Beispiele homogener Typen. Es versteht sich von selbst, daß eine homogene Menge weder ein erstes noch letztes Element noch benachbarte Elemente hat, also offen und dicht ist. Allgemeiner wollen wir noch eine Menge aus mehr als zwei Elementen isomer nennen, wenn alle ihre Mittelstrecken ähnlich sind. Eine isomere Menge ist dicht, braucht aber nicht offen zu sein. Jede Mittelstrecke einer isomeren Menge ist homogen. Isomere Typen sind z. B. $1 + \eta$, $1 + \eta + 1$, $\eta \omega_1$, $(1 + \lambda)\omega_1$ u. dgl.

Wir wollen unsere Betrachtung nur unter verschiedenen spezialisierenden Voraussetzungen durchführen. Zunächst beschränken wir uns auf Potenzen

$$A = (M, m)_{\xi}^{J^*};$$

der Typus μ der Basis M erfahre durch irgend ein Element die Zerlegung

$$\mu = \pi + 1 + \varrho,$$

[1] Hat das Argument ein letztes Element l, so ist das Produkt A nach dem assoziativen Gesetz mit $A_l \cdot A'$ ähnlich, wo A_l der dem Index l entsprechende Faktor ist, hat also Stücke, die mit A_l ähnlich sind, und kann, falls A_l nicht aus einem Element besteht, nur dicht sein, wenn A_l dicht ist.

insbesondere durch das Hauptelement m und durch irgend ein Element $n \neq m$ die Zerlegungen

$$\mu = \pi_m + 1 + \varrho_m, \quad \mu = \pi_n + 1 + \varrho_n.$$

Das Argument J nehmen wir vom Typus einer regulären Anfangszahl ω_σ an und zwar direkt

$$J = W(\omega_\sigma) = \{0, 1, \ldots, i, \ldots\} \qquad (i < \omega_\sigma),$$

so daß die Potenz vom Typus ist

$$\alpha = (\pi_m + 1 + \varrho_m)_\xi^{\omega_\sigma^*}.$$

Wenn dabei $\omega_\sigma < \omega_\xi$ wäre, so würde es sich um eine Vollpotenz handeln, die weniger Interesse bietet; wir setzen daher $\omega_\sigma \geqq \omega_\xi$ voraus.

Da jedes Endstück K von J, falls es nicht verschwindet, wieder vom Typus ω_σ ist[1], so ist für ein solches $\alpha(K) = \alpha$, insbesondere für jede Endstrecke $\alpha(J_i) = \alpha$. Für die durch einen Komplex x von A bewirkte Zerlegung erhalten wir also

$$\alpha = \eta + 1 + \zeta,$$

(1) $$\eta = \sum_i^J \alpha\eta_i = \alpha \sum_i^J \eta_i, \quad \zeta = \sum_i^{J^*} \alpha\zeta_i = \alpha \sum_i^{J^*} \zeta_i,$$

wobei $\mu = \eta_i + 1 + \zeta_i$ die durch x_i bewirkte Zerlegung ist. Insbesondere bewirkt der Hauptkomplex $a = (m, m, \ldots)$ die Zerlegung

$$(\eta_i = \pi_m, \ \zeta_i = \varrho_m)$$

(2) $$\alpha = \alpha\pi_m\omega_\sigma + 1 + \alpha\varrho_m\omega_\sigma^*,$$

die hier allerdings keine „Rekursionsformel" mehr ist (vgl. S. 171).

Unter den sonstigen Zerlegungen betrachten wir zunächst solche durch einen Teilkomplex

$$x_H = (n, n, \ldots),$$

wo $H = W(\omega_\xi)$ und alle $x_h = n \neq m$ sind. Da x_H mit $a_H = (m, m, \ldots)$ nicht für ω_ξ kongruent ist, also kein mit x_H beginnender Komplex in A existiert, während für jedes Anfangsintervall $G = W(h+1)$ noch $x_G \equiv a_G \ (\omega_\xi)$ ist, so bestimmt x_H die Zerlegung analog zu (2)

(3) $$\alpha = \alpha\pi_n\omega_\xi + \alpha\varrho_n\omega_\xi^*.$$

Hieraus folgt z. B.

$$\alpha\pi_n + \alpha = \alpha\pi_n(1 + \omega_\xi) + \alpha\varrho_n\omega_\xi^* = \alpha,$$

dasselbe gilt aber nach (2) auch von π_m. So erhält man für jedes π und ebenso jedes ϱ

(4) $$\alpha(\pi + 1) = \alpha(1 + \varrho) = \alpha.$$

[1] Wir erwähnten schon S. 126, daß diese Eigenschaft nicht allein den Anfangszahlen zukommt.

Überdies ist nach dem assoziativen Gesetz $\alpha = \alpha\mu$; ist ν der Typus irgend einer Mittelstrecke M_p^r, also $\mu = \pi_p + 1 + \nu + 1 + \varrho_r$, so folgt daraus noch

(5) $$\alpha(1 + \nu + 1) = \alpha.$$

α bleibt also nach (4) und (5) bei Multiplikation mit jedem geschlossenen Intervall der Basis (Strecke plus Endpunkten) unverändert.

Ist π von Null verschieden, so kann man setzen $\pi = \pi' + 1 + \nu'$ und findet

$$\alpha\pi = \alpha(\pi' + 1) + \alpha\nu' = \alpha + \alpha\nu'$$

oder für eine beliebige Mittelstrecke ν

$$\alpha\pi = \alpha(1 + \nu + 1) + \alpha\nu' = \alpha(1 + \nu) + \alpha\pi.$$

Es ist also, wenn ν eine beliebige Mittelstrecke, π eine nichtverschwindende Anfangsstrecke, ϱ eine nichtverschwindende Endstrecke bedeutet,

(6) $$\alpha\pi = \alpha(1 + \nu + \pi), \quad \alpha\varrho = \alpha(\varrho + \nu + 1).$$

Ist ferner ω_τ eine reguläre Anfangszahl $< \omega_\xi$, so folgt aus (3) ähnlich wie soeben, wegen $\omega_\tau + \omega_\xi = \omega_\xi$,

$$\alpha\pi_n\omega_\tau + \alpha = \alpha\pi_n(\omega_\tau + \omega_\xi) + \alpha\varrho_n\omega_\xi^* = \alpha.$$

Wir schließen daraus, daß von den niedrigsten Fällen $\omega_\xi = \omega_0$ und $\omega_\xi = \omega_1$ abgesehen, α sicher nicht homogen sein kann, wenn m das erste Element der Basis M ist. Denn dann ist der Hauptkomplex a das erste Element der Potenz A ($\pi_m = 0$, $\alpha = 1 + \alpha\varrho_m\omega_\sigma^*$) und die eben gefundene Zerlegung

$$\alpha = \alpha\pi_n\omega_\tau + 1 + \alpha\varrho_m\omega_\sigma^*$$

zeigt, daß es Elemente von A gibt, deren Anfangsstrecke mit ω_τ konfinal ist[1], also ω_τ-Elemente für jedes reguläre $\omega_\tau < \omega_\xi$. Das widerstreitet für $\omega_\xi > \omega_1$ der Homogenität, die doch gleiche Charaktere aller Elemente fordert. (Nur in den Fällen $\omega_\xi = \omega_0, \omega_1$ ist auch m als erstes oder letztes Element der Basis brauchbar, und man kann mit Hilfe unserer Potenzen oder ihrer Mittelstrecken homogene Typen konstruieren, deren Elemente einen der Charaktere c_{00}, c_{01}, c_{10}, c_{11} haben, wie ich andernorts ausgeführt habe.)

Wir nehmen also jetzt weiter an, daß m weder das erste noch das letzte Element der Basis sei. In (1) sind nun, weil $x = a(\omega_\xi)$ sein muß, alle $x_i = m$, $\eta_i = \pi_m$, $\zeta_i = \varrho_m$ bis auf eine Menge von Stellen, die vom Typus $< \omega_\xi \leqq \omega_\sigma$ ist; insbesondere sind auch bis auf eine solche Ausnahmemenge η_i, ζ_i von Null verschieden und daher η mit ω_σ konfinal, ζ mit ω_σ^* koinitial. Jedes Element

---

[1] Man wende die Anmerkung S. 166 auf ein Produkt an. Hier ist π_n ja sicher von 0 verschieden.

unserer Menge A hat also denselben symmetrischen Charakter $e_{\sigma\sigma}$.

Sodann ergibt sich jetzt durch folgende Schlüsse die Isomerie von A. In (1) bilden die nichtverschwindenden η_i eine Menge vom Typus ω_σ und wir können daher, mit Ersetzung des Argumentes durch ein ähnliches, sämtliche η_i von Null verschieden annehmen. Dann ist nach (6), wenn ν eine beliebige, festgewählte Mittelstrecke von μ bedeutet (man kann z. B. $\nu = 0$ setzen, wenn die Basis benachbarte Elemente hat)

$$\alpha\eta_i = \alpha(1 + \nu + \eta_i).$$

Der Anfang der Summe $\sum_i^J \alpha\eta_i$ ist demnach, weil $\alpha(\eta_i + 1) = \alpha$,

$$\begin{aligned}\sum_i^{W(\omega)} \alpha\eta_i &= \alpha(1 + \nu + \eta_0 + 1 + \nu + \eta_1 + 1 + \nu + \eta_2 + \dots)\\ &= \alpha(1 + \nu + 1 + \nu + 1 + \nu + \dots)\\ &= \alpha(1 + \nu)\omega\end{aligned}$$

und die ganze Summe, die ja nach dem assoziativen Gesetz in lauter solche Teilsummen mit den Argumenten $W(i\omega + \omega) - W(i\omega)$ zerlegt werden kann, ergibt sich so gleich $\alpha(1 + \nu)\omega\omega_\sigma$ oder

$$\sum_i^J \alpha\eta_i = \alpha(1 + \nu)\omega_\sigma$$

(was auch für $\omega_\sigma = \omega$ richtig bleibt). Man sieht, daß hier die η_i eliminiert sind und daß alle Anfangsstrecken und Endstrecken von α die gleichen Typen

$$\beta = \alpha(1 + \nu)\omega_\sigma, \quad \gamma = \alpha(\nu + 1)\omega_\sigma^* \tag{7}$$

haben. Danach sind auch die Mittelstrecken als Endstrecken von Anfangsstrecken (oder umgekehrt) leicht darzustellen. Eine Endstrecke des Produkts $\alpha\alpha'\alpha''$ hat den allgemeinen Typus $\gamma + \alpha\gamma' + \alpha\alpha'\gamma''$, wo $\gamma, \gamma', \gamma''$ Endstrecken von $\alpha, \alpha', \alpha''$ sind.[1] Wendet man dies auf β an, dessen Faktoren $\alpha, 1 + \nu, \omega_\sigma$ die Endstrecken $\gamma, \nu', \omega_\sigma$ haben, so ergibt sich als Typus der Mittelstrecken von A

$$\begin{aligned}&\gamma + \alpha\nu' + \alpha(1 + \nu)\omega_\sigma\\ &= \gamma + \alpha\nu' + \beta\\ &= \gamma + \alpha(\nu + 1) + \alpha\nu' + \alpha(1 + \nu) + \beta\\ &= \gamma + \alpha(\nu + 1 + \nu' + 1 + \nu) + \beta\\ &= \gamma + \alpha(\nu + 1 + \nu) + \beta\\ &= \gamma + \alpha\nu + \beta,\end{aligned}$$

womit also auch das variable ν' noch eliminiert ist. α ist also ein

[1] Vgl. Kap. IV, § 3; natürlich sind das Spezialfälle der in diesem Kapitel, § 6 gegebenen Formeln.

isomerer, sein Mittelstreckentypus $\gamma + \alpha\nu + \beta$ ein homogener Typus, der mit ω_σ konfinal, mit ω_σ^* koinitial ist und lauter $c_{\sigma\sigma}$-Elemente hat.

Auf Homogeneität von α selbst ist im allgemeinen nicht zu rechnen. Wenn die Basis kein letztes Element hat, so ist $\alpha = \alpha\mu$ mit μ konfinal oder mit der regulären Anfangszahl konfinal, mit der μ konfinal ist; hat sie ein letztes Element $n \neq m$, so ist nach (3) $\alpha = \alpha\pi_n\omega_\xi$ mit ω_ξ konfinal. Ebenso ist α entweder mit μ oder mit ω_ξ^* koinitial. Zur Homogeneität von α ist erforderlich, daß es mit ω_σ konfinal und mit ω_σ^* koinitial ist.

Wenn die Basis ein erstes und letztes Element hat (die beide vom Hauptelement verschieden sind), so kann man unter dem ν der letzten Formeln insbesondere die von diesen beiden eingeschlossene Mittelstrecke verstehen, also $\mu = 1 + \nu + 1$ setzen. Aus (3) erhält man, wenn man n mit dem letzten oder ersten Element identifiziert,

$$\alpha = \alpha(1 + \nu)\omega_\xi = \alpha(\nu + 1)\omega_\xi^*.$$

Ist dann insbesondere noch $\omega_\sigma = \omega_\xi$, so wird nach (7) $\alpha = \beta = \gamma$, also α homogen.

Bleiben wir bei dem isomeren α, das mit ω_δ konfinal und mit ω_ε^* koinitial sei, wo die regulären Anfangszahlen ω_δ, ω_ε wie oben zu bestimmen sind. Wir müssen noch die Lückencharaktere von α ermitteln. Alle Zerlegungen

$$\alpha = \eta + \zeta$$

in zwei nichtverschwindende Summanden sind nach S. 170 von der Form

$$\eta = \sum_h^H \alpha\,\eta_h + \bar\eta, \quad \zeta = \bar\zeta + \sum_h^{H^*} \alpha\,\zeta_h,$$

wobei einer der vier Fälle möglich ist:

(a) $\bar\eta = 0,\quad \bar\zeta = 0;$

(b) $\bar\eta = 0,\quad \bar\zeta = \alpha$ oder $1;$

(c) $\bar\eta = \alpha$ oder $1,\quad \bar\zeta = 0;$

(d) $\bar\eta = \alpha\,\eta_k,\quad \bar\zeta = \alpha\,\zeta_k \qquad (\eta_k + \zeta_k = \mu).$

Der Fall (a) gibt, wie wir sahen, eine $c_{\xi\xi}$-Lücke (S. 171 oben).

Der Fall (b) gibt für $H = J$, $\bar\zeta = 1$ die Zerlegung durch ein Element, die wir jetzt beiseite lassen. Für $H < J$, $\bar\zeta = \alpha$ ist H mit einer regulären Anfangszahl $\omega_\tau < \omega_\sigma$ konfinal und mit derselben ist η konfinal. Da ζ dann mit α und mit ω_ε^* koinitial ist, so gibt dies eine $c_{\tau\varepsilon}$-Lücke ($\tau < \sigma$).

Der Fall (c) gibt ebenso eine $c_{\delta\tau}$-Lücke.

Daß diese Lücken, die auftreten können, auch wirklich auftreten, wird z. B. durch die Formeln

$$\alpha = \alpha\pi_n\omega_\xi + \alpha\varrho_n\omega_\xi^* = \alpha\pi_m\omega_\tau + \alpha = \alpha + \alpha\varrho_m\omega_\tau^*$$

angezeigt (die beiden letzten aus (2) folgend). Die erste gibt allerdings nur eine $c_{\xi\xi}$-Lücke, wenn π_n und ϱ_n von 0 verschieden sind, und versagt also in dem einen Fall, daß die Basis aus nur drei Elementen besteht, mit dem mittleren als Hauptelement. Man kann aber statt des Komplexes

$$x_H = (n, n, \ldots), \quad H = W(\omega_\xi),$$

der zu jener Formel (3) geführt hat, etwa

$$x_H = (m, n, m, n, \ldots, m, n, \ldots)$$

wählen, indem man also $x_{2h} = m$, $x_{2h+1} = n$ setzt (vgl. S. 109), und erhält dann

$$\alpha = \alpha(\pi_m + \pi_n)\omega_\xi + \alpha(\varrho_n + \varrho_m)\omega_\xi^*$$

mit sicher von 0 verschiedenen Gliedern.

Der Fall (d) endlich gibt noch weitere Lücken. η_k kann ein letztes Element haben oder mit einer regulären Anfangszahl ω_φ konfinal sein, ζ_k ein erstes Element haben oder mit ω_ψ^* koinitial sein. Man sieht unmittelbar, daß den Charakteren

$$(1, 1), \quad (\omega_\varphi, 1), \quad (1, \omega_\psi^*), \quad (\omega_\varphi, \omega_\psi^*)$$

der Zerlegung $\mu = \eta_k + \zeta_k$ die Charaktere

$$(\omega_\delta, \omega_\varepsilon^*), \quad (\omega_\varphi, \omega_\varepsilon^*), \quad (\omega_\delta, \omega_\psi^*), \quad (\omega_\varphi, \omega_\psi^*)$$

der Zerlegung $\alpha = \eta + \zeta$ entsprechen, also Lücken mit den Charakteren $c_{\delta\varepsilon}$, $c_{\varphi\varepsilon}$, $c_{\delta\psi}$, $c_{\varphi\psi}$. Dies gibt die Möglichkeit, beliebig vorgeschriebene Lücken in die Potenz hineinzubringen dadurch, daß man sie der Basis einpflanzt, was z. B. durch Aufnahme eines Summanden $\omega_\varphi + \omega_\psi^*$ geschehen kann.

Fassen wir die erhaltenen Resultate zusammen. Die Potenz mit der Basis μ

$$\alpha = (\pi + 1 + \varrho)_\xi^{\omega_\sigma^*}, \quad \mu = \pi + 1 + \varrho,$$

worin π, ϱ von Null verschieden und $\omega_\sigma \geqq \omega_\xi$ (beides reguläre Anfangszahlen) angenommen ist, ist i s o m e r mit $c_{\sigma\sigma}$-Elementen. Sie ist mit ω_δ konfinal und mit ω_ε^* koinitial; hierbei ist $\omega_\delta = \omega_\xi$, wenn die Basis μ ein letztes Element hat, andernfalls diejenige reguläre Anfangszahl, mit der μ konfinal ist; ebenso $\omega_\varepsilon = \omega_\xi$, wenn μ ein erstes Element hat, andernfalls diejenige reguläre Anfangszahl, mit der μ^* konfinal ist. Die Potenz α enthält sicher Lücken mit den Charakteren $c_{\xi\xi}$, $c_{\tau\varepsilon}$, $c_{\delta\tau}$, wenn ω_τ eine beliebige reguläre Anfangszahl $< \omega_\sigma$ ist; außerdem Lücken mit den Charakteren $c_{\delta\varepsilon}$, $c_{\varphi\varepsilon}$, $c_{\delta\psi}$, $c_{\varphi\psi}$, wenn die Basis resp. benachbarte Elemente, ω_φ-Elemente, ω_ψ^*-Elemente, $c_{\varphi\psi}$-Lücken besitzt.

Der denkbar einfachste Fall ist nun wohl der, daß die Basis

nicht mehr als die notwendige Anzahl von drei Elementen hat und daß $\omega_\sigma = \omega_\xi$. Wir erhalten so die Potenz

$$\alpha = (1 + 1 + 1)_{\xi}^{\omega_\xi^*}.$$

Hier ist $\omega_\sigma = \omega_\delta = \omega_\varepsilon = \omega_\xi$; α wird jetzt, wie wir sahen, homogen mit lauter $c_{\xi\xi}$-Elementen. Die Lückencharaktere sind $c_{\xi\xi}$, $c_{\tau\xi}$, $c_{\xi\tau}$ für jedes reguläre $\omega_\tau < \omega_\xi$; alle andern Möglichkeiten fallen fort. Wir wollen diesen Typus den der regulären Anfangszahl ω_ξ entsprechenden homogenen Normaltypus nennen und mit

(8) $$\eta_\xi = (1 + 1 + 1)_{\xi}^{\omega_\xi^*}$$

bezeichnen. Man erhält ihn, indem man aus drei Elementen $l < m < n$ diejenigen Elementkomplexe

$$x = (x_0, x_1, \ldots, x_i, \ldots) \qquad (i < \omega_\xi)$$

mit dem Argument $W(\omega_\xi)$ bildet, wo von einem Index an schließlich durchweg $x_i = m$, und diese Komplexe lexikographisch ordnet.

Die Mächtigkeit dieser Normaltypen ist nicht schwer zu bestimmen (ähnliche Betrachtungen sind übrigens auch im allgemeinen Falle anwendbar). Es sei h eine Ordnungszahl $< \omega_\xi$ und $\mathfrak{h}$ ihre Mächtigkeit oder die der Menge $W(h)$. Bilden wir alle Komplexe, für die

$$x_g = l \text{ oder } n \text{ für } g < h, \quad x_i = m \text{ für } i \geqq h,$$

so gehören diese alle unserer Potenz an; ihre Menge hat die Mächtigkeit $2^{\mathfrak{h}}$ und zwei solche Mengen, zu verschiedenen h gehörig, sind fremd. Demnach gilt, die Summe über $W(\omega_\xi)$ erstreckt, für die Mächtigkeit $\mathfrak{y}_\xi$ von η_ξ jedenfalls

$$\mathfrak{y}_\xi \geqq \sum_h 2^{\mathfrak{h}}.$$

Andererseits, bilden wir die Komplexe mit

$$x_g = l, m, n \text{ für } g < h, \quad x_i = m \text{ für } i \geqq h,$$

so hat deren Menge die Mächtigkeit $3^{\mathfrak{h}}$, und die Summe aller dieser Mengen, die aber nicht mehr paarweise fremd sind, gibt alle Komplexe der Potenz. Demnach ist

$$\mathfrak{y}_\xi \leqq \sum_h 3^{\mathfrak{h}}.$$

Für $\xi = 0$ ist

$$\mathfrak{y}_0 \geqq 1 + 2 + 2^2 + 2^3 + \ldots = \aleph_0,$$
$$\mathfrak{y}_0 \leqq 1 + 3 + 3^2 + 3^3 + \ldots = \aleph_0;$$

also $\mathfrak{y}_0 = \aleph_0$. Da η_0 hiernach ein abzählbarer, dichter, offener Typus ist, so ist η_0 mit dem Typus η der Menge der rationalen Zahlen identisch.

Für unendliches $\mathfrak{h}$ ist

$$2^{\mathfrak{h}} = 2^{\aleph_0 \mathfrak{h}} = (2^{\aleph_0})^{\mathfrak{h}} = (3^{\aleph_0})^{\mathfrak{h}} = 3^{\aleph_0 \mathfrak{h}} = 3^{\mathfrak{h}},$$

und da, wie wir soeben sahen, auch die über endliche $\mathfrak{h}$ erstreckten Summen $\Sigma 2^{\mathfrak{h}}$, $\Sigma 3^{\mathfrak{h}}$ übereinstimmen, so ist stets

(9) $$\mathfrak{y}_{\xi} = \sum_h 2^{\mathfrak{h}} \qquad (h < \omega_{\xi})$$

Wir drücken für den Fall, daß $\omega_{\xi} > \omega_0$ keine Limeszahl als Index hat[1], $\mathfrak{y}_{\xi}$ noch etwas einfacher aus. Schreiben wir dann $\xi + 1$ für ξ, so ist

$$\mathfrak{y}_{\xi+1} = \sum_h 2^{\mathfrak{h}} \qquad (h < \omega_{\xi+1}).$$

Beachten wir nur das eine Glied $h = \omega_{\xi}$, so ist

$$\mathfrak{y}_{\xi+1} \geqq 2^{\aleph_{\xi}}.$$

Andererseits ist, für jedes $h < \omega_{\xi+1}$, $\mathfrak{h} < \aleph_{\xi+1}$, also $\mathfrak{h} \leqq \aleph_{\xi}$, demnach

$$\mathfrak{y}_{\xi+1} \leqq \sum_h 2^{\aleph_{\xi}} = \aleph_{\xi+1} \cdot 2^{\aleph_{\xi}}.$$

Da weiter $2^{\aleph_{\xi}} > \aleph_{\xi}$, also $\geqq \aleph_{\xi+1}$, so ist

$$\mathfrak{y}_{\xi+1} \leqq 2^{\aleph_{\xi}} \cdot 2^{\aleph_{\xi}} = 2^{\aleph_{\xi}},$$

also

(10) $$\mathfrak{y}_{\xi+1} = 2^{\aleph_{\xi}}.$$

Wir haben also die Normaltypen

$$\eta_0, \eta_1, \eta_2, \ldots, \eta_{\omega+1}, \ldots$$

mit den Mächtigkeiten

$$\aleph_0, 2^{\aleph_0}, 2^{\aleph_1}, \ldots, 2^{\aleph_{\omega}}, \ldots$$

§ 8. Normaltypen.

Wir haben uns noch klar zu machen, inwiefern die Normaltypen als natürliche Übertragung des Typus η (der Menge der rationalen Zahlen) auf höhere Mächtigkeit anzusehen sind. Es gelten nämlich von ihnen Sätze analog denen in Kap. IV, § 7, insbesondere ein Satz, der aussagt, daß unter gewissen Bedingungen eine Menge sicher vom Normaltypus ist.

Wir verallgemeinern oder verschärfen zunächst den Begriff der offenen dichten Menge in geeigneter Weise. Daß eine solche kein letztes und kein erstes Element und keine benachbarten Elemente hat, können wir auch so ausdrücken: sie ist mit keiner endlichen Teilmenge konfinal oder koinitial und hat kein Paar benachbarter endlicher Teilmengen. Ist nun $\aleph_{\xi}$ irgend ein (nicht notwendig reguläres, d. h. zu einer regulären Anfangszahl ω_{ξ} gehöriges) Alef, so können wir die Forderung stellen, daß eine Menge mit keiner Teilmenge von einer Mächtigkeit $< \aleph_{\xi}$ konfinal

[1] Über etwaige reguläre Anfangszahlen mit Limesindex vgl. S. 131.

oder koinitial sein und kein Paar benachbarter Teilmengen, die beide $< \aleph_\xi$ sind, enthalten soll: eine solche Menge heiße eine η_ξ-Menge. Die offenen dichten Mengen sind hiernach als η_0-Mengen oder η-Mengen zu bezeichnen. Wie man sieht, stellt unsere Forderung eine mit wachsendem ξ zunehmende Verschärfung vor: eine η_ξ-Menge ist zugleich η_τ-Menge für jedes $\tau < \xi$, z. B. ist eine η_2-Menge eine spezielle η_1-Menge und diese wieder eine spezielle η_0-Menge. Es folgt übrigens aus der Definition, daß für ein singuläres $\aleph_\xi$ eine η_ξ-Menge zugleich eine $\eta_{\xi+1}$-Menge ist; denn eine Menge, mit der sie konfinal ist, kann dann auch nicht von der Mächtigkeit $\aleph_\xi$ sein, da diese dann wieder mit einer wohlgeordneten Menge $\leqq \omega_\xi$, also wegen des singulären Charakters von ω_ξ mit einer Menge $< \omega_\xi$ und demnach $< \aleph_\xi$ konfinal wäre (und ebenso für koinitiale und benachbarte Teilmengen). Wir beschränken uns demgemäß wieder auf reguläres $\aleph_\xi$.

Bei allen Zerlegungen einer η_ξ-Menge muß von den beiden regulären Zahlen ϱ, σ, die ihren Charakter (ϱ, σ^*) bilden, mindestens eine von der Mächtigkeit $\geqq \aleph_\xi$ sein. Da ein Elementcharakter $c_{\alpha\beta}$ zwei Zerlegungen mit den Charakteren $(\omega_\alpha, 1)$ und $(1, \omega_\beta{}^*)$ bedingt, so müssen beide Indices (α und β) $\geqq \xi$ sein; bei einem Lückencharakter $c_{\gamma\delta}$ muß mindestens einer der Indices (γ oder δ) $\geqq \xi$ sein. Sind umgekehrt diese Bedingungen erfüllt, so ist vielleicht nicht die Menge selbst (da ja noch die Konfinalitäts- und Koinitialitätsbedingung zu erfüllen bleibt), sicher aber jede ihrer Mittelstrecken eine η_ξ-Menge.

Bei dem am Schlusse von § 7 konstruierten Normaltypus η_ξ ist jeder Elementcharakter $c_{\xi\xi}$, die Lückencharaktere sind

$$c_{0\xi},\ c_{\xi 0},\ c_{1\xi},\ c_{\xi 1},\ \ldots,\ c_{\tau\xi},\ c_{\xi\tau},\ \ldots,\ c_{\xi\xi} \qquad (\tau < \xi);$$

danach ist η_ξ ein η_ξ-Typus.

Es gelten nun folgende beiden Sätze, die mit denen in Kap. IV, § 7 zu vergleichen sind:

I. Eine η_ξ-Menge enthält zu jeder Menge von der Mächtigkeit $\aleph_\xi$ eine ähnliche Teilmenge.

II. Zwei η_ξ-Mengen von der Mächtigkeit $\aleph_\xi$ sind ähnlich.

Bevor wir sie beweisen, bemerken wir dazu: es ist natürlich evident, daß eine η_ξ-Menge mindestens von der Mächtigkeit $\aleph_\xi$ sein muß. Ob es η_ξ-Mengen von der niedrigsten Mächtigkeit $\aleph_\xi$ gibt, ist für $\xi > 0$ noch vollkommen unbekannt (einen η_0-Typus von der Mächtigkeit $\aleph_0$ gibt es, nämlich $\eta_0 = \eta$); der Satz II sagt nur, daß, wenn es welche gibt, sie alle einen und denselben Typus haben müssen. Es wird sich dann weiter herausstellen, daß dieser Typus

der Normaltypus η_ξ sein muß; da nun $\eta_{\xi+1}$ die Mächtigkeit $2^{\aleph_\xi}$ hat, so gibt es $\eta_{\xi+1}$-Mengen von der Mächtigkeit $\aleph_{\xi+1}$ dann und nur dann, wenn $2^{\aleph_\xi} = \aleph_{\xi+1}$. Diese Gleichung ist aber durchaus problematisch, ihr niedrigster Fall $2^{\aleph_0} = \aleph_1$ ist die Cantorsche Kontinuumhypothese.

Zu I ist noch zu bemerken: daß überhaupt eine Menge existiert, die Teilmengen aller Typen von der Mächtigkeit $\aleph_\xi$ enthält, ist an sich nichts Wunderbares, da man zu diesem Zweck ja nur die Summe aller Typen der Typenklasse $T(\aleph_\xi)$ zu bilden braucht. Wichtig ist nur, daß man schon mit einer η_ξ-Menge, z. B. mit dem Normaltypus η_ξ und dessen Mächtigkeit ausreicht (jene Summe würde, wie wir beiläufig ohne Beweis bemerken, die Mächtigkeit $2^{\aleph_\xi}$ von $\eta_{\xi+1}$ haben), und ferner, wenn man sich der Definition von η_ξ erinnert, daß man jede geordnete Menge als lexikographisch geordnete Menge von Komplexen mit wohlgeordnetem Argument und wohlgeordneten Faktoren darstellen kann (sogar schon als Teilmenge einer Potenz mit der Basis 3 und einer regulären Anfangszahl als Argument).

Der Beweis der Sätze I und II verläuft dem in Kap. IV, § 7 so völlig analog, daß wir uns kurz fassen können. A sei eine beliebig geordnete Menge von der Mächtigkeit $\aleph_\xi$, also von der Ordnung abgesehen in der Form

$$A = \{a_0, a_1, \ldots, a_\alpha, \ldots\} \qquad (\alpha < \omega_\xi)$$

darstellbar, und B eine η_ξ-Menge. Wir ordnen dem a_0 ein beliebiges Element b_0 von B zu und zeigen induktiv, daß man unter Erhaltung der Ordnung jedem a_α ein b_α zuordnen kann, wenn dies für alle Zahlen $< \alpha$ bereits geschehen. Die Menge $A_\alpha = \{a_0, a_1, \ldots, a_\lambda, \ldots\}$ $(\lambda < \alpha)$ wird durch das Element a_α in zwei Stücke $A_\alpha = A_\alpha' + A_\alpha''$ zerlegt; A_α' ist die Menge der Elemente von A_α, die vor a_α liegen, A_α'' die Menge derer, die nach a_α folgen. Eins dieser Stücke kann 0 sein. Das Bild B_α zerfällt entsprechend in $B_\alpha' + B_\alpha''$, und das Bild b_α des Elementes a_α ist so zu wählen, daß es zwischen B_α' und B_α'' fällt, eventuell der ganzen Menge B_α vorangeht oder nachfolgt. Ein solches Element ist aber stets in B vorhanden, da die Mengen $B_\alpha, B_\alpha', B_\alpha''$ von einer Mächtigkeit $< \aleph_\xi$ sind, also B mit B_α weder konfinal noch koinitial ist und B_α', B_α'' auch nicht benachbart sind. Damit ist I bewiesen.

Sind ferner

$$A = \{a_0, a_1, \ldots, a_\alpha, \ldots\} \qquad (\alpha < \omega_\xi)$$

$$B = \{b_0, b_1, \ldots, b_\beta, \ldots\} \qquad (\beta < \omega_\xi)$$

zwei η_ξ-Mengen von der Mächtigkeit $\aleph_\xi$, so wenden wir wieder

alternierende Abbildung an. Wir ordnen jeder Ordnungszahl $\gamma < \omega_\xi$ ein Paar a^γ, b^γ zu, das wir induktiv durch die Mengen

$$A^\gamma = \{a^0, a^1, \ldots, a^\lambda, \ldots\} \qquad (\lambda < \gamma)$$
$$B^\gamma = \{b^0, b^1, \ldots, b^\lambda, \ldots\}$$

der bereits gepaarten Elemente so definieren: $a^{2\gamma}$ sei in der Menge $A - A^{2\gamma}$ das Element a_α mit niedrigstem Index und $b^{2\gamma}$ ein Element, das zu den Elementen von $B^{2\gamma}$ dieselbe Ordnung hat wie $a^{2\gamma}$ zu den Elementen von $A^{2\gamma}$; umgekehrt sei $b^{2\gamma+1}$ in der Menge $B - B^{2\gamma+1}$ das Element b_β mit niedrigstem Index und $a^{2\gamma+1}$ ein wie soeben definiertes passendes Bild. (Den Anfang macht $a^0 = a_0$ und ein beliebiges Element b^0). Die Existenz solcher Bilder ist gesichert, weil A und B η_ξ-Mengen sind, und die Definition von $a^{2\gamma}$ und $b^{2\gamma+1}$ schließt das Übergehen eines Elementes aus; man kann beiläufig leicht zeigen, daß a_α in $A^{2\alpha+1}$ und b_β in $B^{2\beta+2}$ vorkommen muß. Also wird A auf B ähnlich abgebildet und II ist bewiesen.

Man kann noch die Abbildung (bei gegebener Wohlordnung von A und B in der obigen Gestalt) dadurch vollkommen willkürfrei machen, daß man auch für die Bilder $b^{2\gamma}, a^{2\gamma+1}$ die verfügbaren Elemente mit kleinstem Index wählt; das gleiche ist beim Beweise des Satzes I durch eine präliminarische Wohlordnung von B zu erreichen.

Weiter gilt noch der Satz:

III. Jede η_ξ-Menge enthält eine Teilmenge vom Normaltypus η_ξ.

Dies ist für $\xi = 0$ eine bloße Folge von I, da η_0 abzählbar ist; für $\xi > 0$, wo wir nicht wissen, ob η_ξ die Mächtigkeit $\aleph_\xi$ hat, ist es eine wesentliche Ergänzung, auf Grund deren wir sofort behaupten können: der Normaltypus η_ξ hat als η_ξ-Typus die kleinste denkbare Mächtigkeit, und wenn es einen η_ξ-Typus der Mächtigkeit $\aleph_\xi$, also nach II nur einen einzigen, gibt, so ist dies notwendig der Normaltypus η_ξ.

A sei eine Menge vom Normaltypus η_ξ und B eine η_ξ-Menge. Wir erinnern uns der Darstellung von A (§ 7) als Menge der Komplexe

$$x = (x_0, x_1, \ldots, x_\alpha, \ldots) \qquad (\alpha < \omega_\xi)$$

mit dem Argument $W(\omega_\xi)$, wobei jedes x_α eins von drei Elementen $l < m < n$ und von einem gewissen Index an durchweg $x_\alpha = m$ ist. Die kleinste Ordnungszahl α, die die Eigenschaft hat, daß, für $\beta \geqq \alpha$, $x_\beta = m$ ist, werde mit $\alpha = \varphi(x)$ bezeichnet, und die Menge derjenigen Komplexe von A, für die $\varphi(x) = \alpha$, mit A_α; dann ist

$$A = \sum_\alpha A_\alpha = A_0 + A_1 + \ldots + A_\alpha + \ldots \qquad (\alpha < \omega_\xi),$$

welche Formel im Sinne der Addition ungeordneter Mengen zu ver-

stehen ist. Die Menge A_0 besteht aus dem einen Komplex $(m, m, m, \ldots)$, die Menge A_1 aus den beiden $(l, m, m, \ldots)$ und $(n, m, m, \ldots)$, die Menge A_2 aus den sechs Komplexen, die mit ll, ln, ml, mn, nl, nn beginnen usw. Die Menge

$$A^\alpha = A_0 + A_1 + \ldots + A_\alpha$$

besteht dann aus allen Komplexen, für die $x_\beta = m$ für $\beta \geqq \alpha$, während die den Indices vor α entsprechenden Elemente alle drei Werte l, m, n durchlaufen können; die auf das Argument $W(\alpha)$ beschränkten Anfangskomplexe dieser Komplexe durchlaufen also die Vollpotenz mit der Basis $\{l, m, n\}$ und dem Argument $W(\alpha)$, und es ist demnach A^α vom Typus $3^{\alpha *}$ (für $\alpha = 0$ vom Typus 1). Auf Grund des Satzes in § 5 (S. 167) schließen wir daraus sofort, daß jede (nichtverschwindende) Teilmenge von A^α mit einer Menge von der Mächtigkeit $< \aleph_\xi$ konfinal und mit einer ebensolchen koinitial ist. Denn sie ist mit einer Menge konfinal, die mit einer Teilmenge eines Faktors oder des Arguments oder des Exponenten ähnlich ist; die Faktoren sind vom Typus 3, der Exponent vom Typus α^*, und deren Teilmengen (> 0) haben stets ein letztes Element. Unsere Menge ist also entweder mit 1 oder mit einer Ordnungszahl $\leqq \alpha < \omega_\xi$ konfinal, und damit ist die Behauptung bewiesen.

Nunmehr ist der Beweis von III genau wie der von I zu führen. Wir denken uns die einzelnen Mengen A_α wohlgeordnet und damit auch eine Wohlordnung der Menge

$$A = A_0 + A_1 + \ldots + A_\alpha + \ldots$$

herbeigeführt, in der die Summanden in der Reihenfolge ihrer Indices stehen. Wir zeigen induktiv, daß man unter Erhaltung der Rangordnung jedem Komplex x von A ein Element von B zuordnen kann, wenn dies für die (in der Wohlordnung) vorangehenden Komplexe bereits geschehen ist.[1] Deren Menge $A(x)$ wird durch x in zwei Summanden $A'(x) + A''(x)$ zerlegt (mit Rücksicht auf die Rangordnung in der geordneten Menge A), deren einer auch verschwinden kann, und wenn $B(x)$, $B'(x)$, $B''(x)$ die entsprechenden Bilder sind, so hat man als Bild von x ein Element von B zu suchen, das zwischen $B'(x)$ und $B''(x)$ liegt, resp. im Falle des Verschwindens eines der Summanden ein Element vor oder nach $B(x)$. Die Existenz solcher Elemente ist aber sicher. Denn gehört x etwa dem Summanden A_α an, so sind $B(x)$, $B'(x)$, $B''(x)$ Bilder gewisser Teilmengen von A^α, also im Fall des Nichtverschwindens mit Mengen $< \aleph_\xi$ konfinal und koinitial, so daß B als η_ξ-Menge mit $B(x)$ weder konfinal noch koinitial

[1] Den Anfang macht der eine Komplex von A_0, dem ein beliebiges Element von B zuzuordnen ist.

ist und die Mengen $B'(x)$, $B''(x)$ auch nicht benachbart sein können. A läßt sich also auf eine Teilmenge von B ähnlich abbilden.

Die Betrachtungen dieses Paragraphen zeigen, daß wirklich die Normaltypen für höhere Mächtigkeiten genau dieselbe Rolle spielen wie der Typus η der Menge der rationalen Zahlen für die Mächtigkeit $\aleph_0$, soweit man bei den ungelösten Mächtigkeitsproblemen eine solche Übertragung überhaupt wagen kann.

§ 9. **Rationale Ordnungszahlen.**

Anhangsweise wollen wir noch eine vollkommen andersgeartete Verallgemeinerung der Menge der positiven rationalen Zahlen auf höhere Mächtigkeit zur Sprache bringen. Verstehen wir unter α, β zwei Ordnungszahlen > 0, so liegt es nahe, daraus ein geordnetes Paar $\alpha | \beta$ zu bilden, worin α der Zähler, β der Nenner heißen möge, und diesen Paaren in ähnlicher Weise wie den Paaren natürlicher Zahlen durch geeignete Definitionen von Gleichheit und Verschiedenheit, Größenordnung und Verknüpfungen (Addition usw.) den Charakter von gebrochenen oder rationalen Ordnungszahlen aufzuprägen, unter denen die Cantorschen Ordnungszahlen als ganze rationale Ordnungszahlen mit dem Nenner 1 einen Spezialfall bilden. Nun hat die Definition geeigneter Verknüpfungen allerdings ihre Schwierigkeiten, aber Gleichheit, Verschiedenheit und Rangordnung lassen sich zwanglos erklären, wenn man sich des Euklidischen Algorithmus für Ordnungszahlen (Kap. V, § 2) oder der Kettenbruchentwicklung erinnert. Beschränken wir uns auf Paare, wo der Zähler größer ist als der Nenner, und schreiben dafür $\alpha | \alpha_1$ $(\alpha > \alpha_1)$. Die Zahlen α, α_1 bestimmen eindeutig das Formelsystem

$$(1) \quad \left\{ \begin{aligned} \alpha &= \alpha_1 \xi_1 + \alpha_2 \\ \alpha_1 &= \alpha_2 \xi_2 + \alpha_3 \\ &\cdot \quad \cdot \quad \cdot \\ \alpha_{m-2} &= \alpha_{m-1} \xi_{m-1} + \alpha_m \\ \alpha_{m-1} &= \alpha_m \xi_m \end{aligned} \right.$$

worin m eine natürliche Zahl,

$$(2) \qquad \alpha > \alpha_1 > \alpha_2 > \dots > \alpha_{m-1} > \alpha_m \geqq 1,$$

$$(3) \qquad \xi_1, \xi_2, \dots, \xi_{m-1} \geqq 1, \quad \xi_m > 1.$$

D. h. sie bestimmen den „Quotienten“ ξ_1 und den „Rest“ α_2 der „Division von α durch α_1“; dann den Quotienten ξ_2 und den Rest α_3 der Division von α_1 durch α_2 usf. bis zum letzten nicht verschwindenden Rest α_m, für den die letzte Division von α_{m-1} durch α_m „aufgeht“. Sie bestimmen also den Komplex der Quotienten

$$(4) \qquad \xi = (\xi_1, \xi_2, \dots, \xi_m),$$

dessen Elemente Ordnungszahlen $\geqq 1$, die letzte $\xi_m > 1$, sind, und wir werden in erster Linie zwei Zahlenpaare $\alpha|\alpha_1$ und $\beta|\beta_1$ als gleich definieren, wenn sie denselben Quotientenkomplex bestimmen.[1]

Man sieht nun, daß zu jedem Komplex ξ mit den Bedingungen (3) verschiedene Paare $\alpha|\alpha_1$ gehören. Multipliziert man z. B. die Gleichungen (1) links mit einer beliebigen Ordnungszahl $\mu > 0$, so ist

$$\mu\alpha = \mu(\alpha_1\xi_1 + \alpha_2) = \mu\alpha_1 \cdot \xi_1 + \mu\alpha_2$$

und $\mu\alpha > \mu\alpha_1 > \mu\alpha_2$, woraus sich durch Fortsetzung ergibt, daß das Paar $\mu\alpha|\mu\alpha_1$ denselben Komplex ξ hat wie das Paar $\alpha|\alpha_1$ und ihm also gleichzusetzen ist; man darf Zähler und Nenner links mit derselben Zahl multiplizieren, den „Bruch erweitern". Andererseits sieht man durch Rückwärtslesen von (1), daß durch Angabe des Quotientenkomplexes ξ und des letzten Restes α_m die Zahlen $\alpha_{m-1}, \ldots, \alpha_2, \alpha_1, \alpha$ bestimmt sind.[2] Gleiche Zahlenpaare, d. h. solche mit gleichem ξ, können sich also nur durch das zugehörige α_m unterscheiden. Unter ihnen heben wir das der Annahme $\alpha_m = 1$ entsprechende reduzierte Paar $\varrho|\varrho_1$ hervor, das also durch die Gleichungen

$$(5) \qquad \begin{cases} \varrho = \varrho_1\xi_1 + \varrho_2 \\ \varrho_1 = \varrho_2\xi_2 + \varrho_3 \\ \quad \cdot \quad \cdot \quad \cdot \\ \varrho_{m-2} = \varrho_{m-1}\xi_{m-1} + 1 \\ \varrho_{m-1} = \quad \xi_m \end{cases}$$

definiert ist; man sieht dann, daß $\alpha = \alpha_m\varrho$, $\alpha_1 = \alpha_m\varrho_1$ ist. Zahlenpaare sind also gleich, wenn sie zu demselben reduzierten Paar gehören, und entstehen aus diesem durch Erweiterung.

Die Analogie mit den natürlichen Rationalzahlen geht aber noch weiter und erlaubt auch eine Größenordnung verschiedener Zahlenpaare zu definieren. Denken wir daran: bei einer gewöhnlichen Kettenbruchentwicklung

$$\frac{\alpha}{\alpha_1} = \xi_1 + \cfrac{1}{\xi_2 + \cfrac{1}{\xi_3 + \cdots}}$$

bewirkt eine Vergrößerung des ersten Quotienten ξ_1 eine Vergrößerung des Bruches; eine Vergrößerung des zweiten Quotienten ξ_2 bewirkt, bei unverändertem ξ_1, eine Verkleinerung des Bruches bis zu dem für $\lim \xi_2 = \infty$ eintretenden Grenzwert ξ_1; eine Vergrößerung des dritten Quotienten, ohne Änderung der beiden ersten, bewirkt wieder

[1] Zwei Komplexe $\xi = (\xi_1, \xi_2, \ldots, \xi_m)$ und $\eta = (\eta_1, \eta_2, \ldots, \eta_n)$ sind natürlich dann und nur dann gleich, wenn $m = n$ und $\xi_1 = \eta_1, \xi_2 = \eta_2, \ldots, \xi_m = \eta_m$.

[2] Der letzte Rest α_m ist, wie sich der Leser klar machen möge, der „größte gemeinsame linke Teiler" von α und α_1.

eine Vergrößerung des Bruches bis zu dem für $\lim \xi_3 = \infty$ eintretenden Grenzwerth $\xi_1 + \frac{1}{\xi_2}$ usw. Wir werden danach zwei verschiedene Komplexe ξ, η in der folgenden Weise ordnen. Verstehen wir unter ∞ ein Element (z. B. eine Ordnungszahl), das allen in ξ und η auftretenden Ordnungszahlen nachfolgt, so ergänzen wir zunächst

$$\xi = (\xi_1, \xi_2, \ldots, \xi_m) = (\xi_1, \xi_2, \ldots, \xi_m, \infty, \infty, \infty, \ldots)$$
$$\eta = (\eta_1, \eta_2, \ldots, \eta_n) = (\eta_1, \eta_2, \ldots, \eta_n, \infty, \infty, \infty, \ldots)$$

zu Komplexen mit der Menge aller natürlichen Zahlen als Argument und definieren dann $\xi < \eta$, wenn

$$\begin{aligned} &\text{entweder} && \xi_1 < \eta_1 \\ &\text{oder} && \xi_1 = \eta_1,\ \xi_2 > \eta_2 \\ &\text{oder} && \xi_1 = \eta_1,\ \xi_2 = \eta_2,\ \xi_3 < \eta_3 \quad \text{usw.} \end{aligned}$$

Daß aus $\xi < \eta$, $\eta < \zeta$ stets $\xi < \zeta$ folgt, ist leicht zu beweisen. Danach haben z. B. die folgenden Komplexe die Rangordnung, in der sie geschrieben sind:

$$(2)\ (2, \omega)\ (2, 3)\ (2, 2)\ (2, 1, 2)\ (2, 1, 3)\ (2, 1, \omega)\ (3) \ldots$$

Irgend eine Menge von Komplexen ξ wird durch diese Festsetzungen zu einer geordneten Menge. Natürlich ist dies ein Spezialfall von lexikographischer Ordnung, woran durch den Umstand nichts geändert wird, daß ξ und η, bei erster Differenzstelle n, für gerades n die umgekehrte Ordnung wie ξ_n und η_n haben. Ist W die Menge der in allen Komplexen ξ auftretenden Elemente $\xi_1, \xi_2, \ldots,$ das Element ∞ mitgerechnet, so ist unsere Komplexmenge Teilmenge des (über die Menge der natürlichen Zahlen als Argument erstreckten) Vollprodukts

$$\mathfrak{P}_i^J A_i = \ldots A_4 A_3 A_2 A_1,$$

wobei

$$A_1 = A_3 = \ldots = W, \ A_2 = A_4 = \ldots = W^*,$$

sogar spezieller des Produkts vom Grade 0

$$\mathfrak{P}_i^J (A_i, \infty)_0$$

mit dem Hauptelement ∞ für jeden Faktor. Sie umfaßt aber niemals alle Komplexe dieses Produkts, sondern nur solche, wo zuerst Nebenelemente $\xi_1, \xi_2, \ldots, \xi_m$ (deren letztes überdies > 1) stehen und dann das Hauptelement ∞ ununterbrochen auftritt, während das Produkt auch Komplexe wie $(\infty, 2, \infty, \infty, \ldots)$ enthält, in denen auf ein Hauptelement noch Nebenelemente folgen.

Bilden wir speziell, für eine reguläre Anfangszahl ω_σ, alle Zahlenpaare $\alpha|\alpha_1$ mit $1 \leqq \alpha_1 < \alpha < \omega_\sigma$ oder alle Komplexe (4), die mit Beachtung der Ungleichungen (3) aus Zahlen von $W(\omega_\sigma)$ gebildet sind, so ist deren geordnete Menge A als Verallgemeinerung der Menge der rationalen Zahlen > 1 anzusehen. Wir können die Haupteigenschaften von A leicht feststellen.

Es gibt keinen Komplex, der alle Komplexe

$$(2)\ (3)\ \ldots\ (\alpha)\ \ldots \qquad (\alpha < \omega_\sigma)$$

überträfe; A ist also mit der Menge dieser Komplexe oder mit ω_σ konfinal. Ebenso ist A mit der Menge der Komplexe

$$(1,2)\ (1,3)\ \ldots\ (1,\alpha)\ \ldots$$

oder mit ω_σ^* koinitial.

Jedes Element ξ ist ω_σ-Element oder ω_σ^*-Element. Denn $\xi = (\xi_1, \xi_2, \ldots, \xi_m)$ ist jedenfalls Limes der Menge der Komplexe

$$(\xi, 2)\ (\xi, 3) \ldots (\xi, \alpha) \ldots,$$

wobei (ξ, α) für $(\xi_1, \xi_2, \ldots, \xi_m, \alpha)$ gesetzt ist. Diese Menge ist für gerades m vom Typus ω_σ und hat ξ zum oberen Limes, für ungerades m ist sie vom Typus ω_σ^* und hat ξ zum unteren Limes.

Ist ξ_m keine Limeszahl, also mit unmittelbarem Vorgänger $\xi_m - 1 \geqq 1$ versehen, so setze man

$$\xi' = (\xi_1, \xi_2, \ldots, \xi_{m-1}, \xi_m - 1),$$

welches übrigens möglicherweise (für $\xi_m = 2$) kein Komplex unserer Menge ist, und betrachte die Menge der Komplexe

$$(\xi', 1, 2)\ (\xi', 1, 3) \ldots (\xi', 1, \alpha) \ldots,$$

die für gerades m vom Typus ω_σ^*, für ungerades m vom Typus ω_σ ist und beide Male ξ zum (unteren resp. oberen) Limes hat. Die Komplexe ξ, wo ξ_m keine Limeszahl ist, sind also Elemente vom Charakter $c_{\sigma\sigma}$.

Ist ξ_m Limeszahl und mit der regulären Anfangszahl $\omega_\tau < \omega_\sigma$ konfinal, so setze man

$$\xi'' = (\xi_1, \xi_2, \ldots, \xi_{m-1})$$

und betrachte die Menge der Komplexe

$$(\xi'', 2)\ (\xi'', 3) \ldots (\xi'', \beta) \ldots \qquad (\beta < \xi_m),$$

die für gerades m vom Typus ξ_m^*, für ungerades vom Typus ξ_m ist und wieder ξ zum Limes hat.[1] Ist also ξ_m Limeszahl und mit ω_τ konfinal, so hat ξ für gerades m den Charakter $c_{\sigma\tau}$, für ungerades den Charakter $c_{\tau\sigma}$.

Die Elementcharaktere von A sind also

$$c_{0\sigma}\, c_{\sigma 0}\, c_{1\sigma}\, c_{\sigma 1} \ldots c_{\tau\sigma}\, c_{\sigma\tau} \ldots c_{\sigma\sigma} \qquad (\tau < \sigma),$$

[1] Für $m = 1$ ist natürlich die Menge der Komplexe (2) (3) … (β) … zu nehmen.

dieselben, die nach S. 181 als Lückencharaktere in dem Normaltypus η_σ auftreten.

Die Lücken von A sind sämtlich vom Charakter c_{00}, wie man durch eine Anwendung der allgemeinen Zerlegungsformeln von § 5 erkennt, die wir dem Leser überlassen wollen.

Man kann den Charakter eines Zahlenpaars $\alpha | \alpha_1$ auch aus der reduzierten Darstellung erkennen. Nach (1) ist α mit α_2, α_2 mit α_4 usw. konfinal, ebenso α_1 mit α_3, α_3 mit α_5 usw. Bei geradem m ist also α mit α_m, α_1 mit α_{m-1} konfinal, bei ungeradem α mit α_{m-1} und α_1 mit α_m. In der reduzierten Darstellung ist $\alpha_m = 1$, $\alpha_{m-1} = \xi_m$, also eine der beiden Zahlen α, α_1 mit 1 und die andere mit ξ_m konfinal. Demnach folgt: ist in der reduzierten Darstellung weder Zähler noch Nenner eine Limeszahl, so hat das Paar den Charakter $c_{\sigma\sigma}$; ist der Zähler eine Limeszahl (mit $\omega_\tau < \omega_\sigma$ konfinal), so hat das Paar den Charakter $c_{\tau\sigma}$; ist der Nenner eine, so hat das Paar den Charakter $c_{\sigma\tau}$. Z. B. haben für $\sigma = 1$ die Paare $3|2$, $\omega|1$, $\omega+1|\omega$ die Charaktere c_{11}, c_{01}, c_{10}.

Die Menge A ist nicht homogen, da sie verschiedene Elementcharaktere besitzt; nichtsdestoweniger hat sie einige der Homogenität ähnliche Eigenschaften, z. B. daß der Typus einer Mittelstrecke nur von den Charakteren der einschließenden Elemente abhängt; wir verzichten darauf einzugehen.

Es ist evident, wie man den Paaren $\alpha | \beta$ mit $\alpha > \beta$ links noch die Paare $\alpha | \alpha = 1 | 1$ und die Paare $\alpha | \beta$ mit $\alpha < \beta$ anzureihen hat, um die Menge A zu einem Analogon der Menge aller positiven Rationalzahlen zu vervollständigen.

§ 10. Initiale und finale Ordnung.

Die lexikographische Ordnung, an die wir uns bisher ausschließlich gehalten haben, entbehrt trotz ihrer natürlichen Einfachheit nicht einer gewissen Willkür und läßt sich durch andere Ordnungsprinzipien ersetzen. So könnte man, um Komplexe $a = (\ldots, a_i, \ldots)$ mit dem Argument J zu ordnen, folgendes verabreden: wenn J ein nichtverschwindendes A n f a n g s s t ü c k H hat, für dessen Elemente h

$$\text{stets} \quad a_h \leqq b_h, \quad \text{aber nicht stets} \quad a_h = b_h$$

ist, so soll $a < b$ gelten, und $a > b$ soll mit $b < a$ gleichbedeutend sein. Man überzeugt sich sofort, daß diese i n i t i a l e O r d n u n g das transitive Gesetz befolgt und die lexikographische als Spezialfall enthält; aber auch sie vermag eine beliebige Komplexmenge im allgemeinen nur teilweise zu ordnen.

Eine Modifikation hiervon, die wir auch noch als initiale Ordnung bezeichnen wollen, die aber nicht mehr als Verallgemeinerung

der lexikographischen Ordnung angesehen werden kann, besteht darin, daß wir in der obigen Forderung das Zeichen $\leqq$ zu $<$ verschärfen. Wenn also ein von Null verschiedenes Anfangsstück H des Arguments existiert, wo durchweg

$$a_h < b_h,$$

so soll $a < b$ sein.

Es bleibt uns dabei unbenommen, den Fall, daß weder $a < b$, noch $a > b$ ist, noch in anderer Weise zu zerlegen, als wir dies bisher getan haben (in $a = b$ und $a \parallel b$, § 1). Für Anwendungszwecke empfiehlt es sich, den Fall auszuzeichnen, daß in einem nichtverschwindenden Anfangsstück H von J durchweg

$$a_h = b_h$$

ist. Wir wollen in diesem Fall, statt ein neues Zeichen einzuführen, geradezu $a = b$ definieren, womit wir also (nur in diesem Paragraphen) dem Gleichheitszeichen zwischen Komplexen einen neuen, erweiterten Sinn beilegen. Die Berechtigung hierzu liegt in der Gültigkeit gewisser formaler Gesetze, die wir nun einmal vom Gleichheitszeichen unter allen Umständen verlangen, und die wir sofort aufstellen werden. Zuvor verabreden wir noch für den vierten Fall, daß weder $a < b$, noch $a = b$, noch $a > b$, das frühere, jetzt in engerem Sinne zu verstehende Zeichen $a \parallel b$, und nennen dann a mit b unvergleichbar, in den drei ersten Fällen a mit b vergleichbar. Wir definieren also

$$(1) \qquad \begin{cases} a < b, & \text{wenn in } H \quad a_h < b_h, \\ a = b, & \text{wenn in } H \quad a_h = b_h, \\ a > b, & \text{wenn in } H \quad a_h > b_h, \\ a \parallel b & \text{in jedem andern Fall.} \end{cases}$$

Dabei ist unter H ein nichtverschwindendes Anfangsstück von J verstanden.

Die formalen Gesetze dieser vier Zeichen (von denen ϱ ein beliebiges bedeute) sind:

$$(2) \qquad a = a.$$

$$(3) \qquad \begin{array}{ll} \text{Aus} & a < b,\ a = b,\ a > b,\ a \parallel b \\ \text{folgt} & b > a,\ b = a,\ b < a,\ b \parallel a. \end{array}$$

$$(4) \qquad \begin{cases} \text{Aus} & a = b,\ b \varrho c & \text{folgt} & a \varrho c, \\ \text{,,} & a \varrho b,\ b = c & \text{,,} & a \varrho c, \\ \text{,,} & a < b,\ b < c & \text{,,} & a < c, \\ \text{,,} & a > b,\ b > c & \text{,,} & a > c. \end{cases}$$

Aus ihnen, deren leichten Beweis wir dem Leser überlassen, geht auch die Zulässigkeit des Gleichheitszeichens hervor.

Statt von der lexikographischen Ordnung (nach ersten Differenzen) hätten wir nun auch von der antilexikographischen nach letzten Differenzen ausgehen und diese zu entsprechenden finalen Ordnungen verallgemeinern oder modifizieren können; offenbar ändert sich dadurch an allem Bisherigen nichts weiter, als daß wir unter H ein nichtverschwindendes Endstück des Arguments J zu verstehen haben. Für Anwendungen ziehen wir das Schema der finalen Ordnung vor, wie es durch die Definitionen (1) erklärt ist ($H > 0$ ein Endstück von J).

Ein Produkt $\mathfrak{P}_i^J A_i$ wird hierdurch wieder zu einer im allgemeinen nur teilweise geordneten Menge. Eine geordnete Teilmenge ist dadurch definiert, daß für je zwei verschiedene (d. h. nicht identische, also nicht im früheren, engeren Sinne gleiche) Komplexe a, b von ihr eine der Relationen $a \lesseqgtr b$ bestehen soll; sie darf also auch von jeder Klasse gleicher Komplexe höchstens ein Element enthalten, und sie darf keine unvergleichbaren Elemente enthalten. Der Begriff einer größten geordneten Teilmenge ist wie in § 1 zu erklären, und man hätte eine Reihe Probleme vor sich wie bei der lexikographischen Ordnung, nämlich besonders interessante geordnete Teilmengen und eventuell größte geordnete Teilmengen auf ihren Typus zu untersuchen.

Wir begnügen uns mit einem einzigen sehr speziellen Beispiel, weil es das Modell abgibt für eine Anzahl Probleme der Analysis, die vielfach, aber ohne definitiven Erfolg behandelt worden sind.

Das Argument J sei die Menge der natürlichen Zahlen, und unsere Komplexe[1]

$$a = (a_1, a_2, \ldots, a_n, \ldots)$$

sollen reelle Zahlen a_n zu Elementen haben, also einfach Folgen reeller Zahlen sein. Bei finaler Ordnung ist also

$$a < b, \; a = b, \; a > b$$

zu definieren, wenn schließlich (d. h. von einem gewissen Index an)

$$a_n < b_n, \; a_n = b_n, \; a_n > b_n$$

ist, und $a \| b$ in jedem andern Falle. Die Menge aller reellen Zahlenfolgen hat, wie wir wissen, die Mächtigkeit des Kontinuums $\aleph^{\aleph_0} = \aleph$.

Hieraus ergibt sich nun ein sehr einfacher Satz:

I. Zu jeder höchstens abzählbaren Menge von Zahlenfolgen gibt es eine größere und eine kleinere Zahlenfolge. Zu zwei höchstens abzählbaren Mengen von Zahlenfolgen,

---

[1] $a \equiv b$ bedeute, daß a und b im engeren Sinne gleich sind, also durchweg $a_n = b_n$.

falls jede Zahlenfolge der einen jeder Zahlenfolge der andern vorangeht, gibt es eine zwischen beiden Mengen liegende Zahlenfolge.

Es sei $F = (a, b, c, \ldots)$ eine Folge von Komplexen oder Zahlenfolgen.[1] Wählen wir dann eine Zahlenfolge x gemäß den Ungleichungen

$$x_1 > a_1,$$
$$x_2 > a_2,\ b_2,$$
$$x_3 > a_3,\ b_3,\ c_3 \text{ usw.}$$

so ist $x > a$, $x > b$, $x > c, \ldots$ Ebenso kann man ein x angeben, das allen Elementen von F vorangeht.

Ferner seien F und $G = (p, q, r, \ldots)$ zwei solche Folgen, und jedes Element von F gehe jedem Element von G voran, also

$$a, b, c, \ldots < p, q, r, \ldots$$

Man kann demgemäß eine Reihe wachsender natürlicher Zahlen $\lambda < \mu < \nu < \pi < \ldots$ so bestimmen, daß

$$a_n < p_n \quad \text{für } n \geqq \lambda,$$
$$a_n,\ b_n < p_n,\ q_n \quad \text{für } n \geqq \mu,$$
$$a_n,\ b_n,\ c_n < p_n,\ q_n,\ r_n \quad \text{für } n \geqq \nu \text{ usw.}$$

Wählt man dann eine Zahlenfolge x gemäß den Ungleichungen

$$a_n < x_n < p_n \quad \text{für } \lambda \leqq n < \mu,$$
$$a_n,\ b_n < x_n < p_n,\ q_n \quad \text{für } \mu \leqq n < \nu,$$
$$a_n,\ b_n,\ c_n < x_n < p_n,\ q_n,\ r_n \quad \text{für } \nu \leqq n < \pi \text{ usw.,}$$

so ist $a < x < p$, $b < x < q$, $c < x < r, \ldots$ oder $a, b, c, \ldots < x < p, q, r, \ldots$, x liegt zwischen den Elementen von F und G.

Für eine größte geordnete Menge von Zahlenfolgen folgt aus I unmittelbar, daß sie mit keiner Menge von der Mächtigkeit $\leqq \aleph_0$ konfinal oder koinitial sein kann und daß in ihr zwei Mengen von der Mächtigkeit $\leqq \aleph_0$ niemals benachbart sein können. Sie ist also (§ 8) eine η_1-Menge, von welchem Begriff wir hier somit eine wichtige Anwendung erhalten. Da sie eine Menge vom Normaltypus η_1 und der Mächtigkeit $2^{\aleph_0} = \aleph$ enthält und ihrerseits in der Menge aller Zahlenfolgen von der Mächtigkeit $\aleph$ enthalten ist, so ist sie genau von der Mächtigkeit des Kontinuums.

Damit ist freilich alles Wesentliche erschöpft, was wir über diese Mengen wissen. Wenn das Kontinuum von der zweiten Mächtigkeit $\aleph_1$ ist, so sind alle solche Mengen vom Normaltypus η_1.

---

[1] Sollte es sich nur um eine endliche Menge handeln, so denke man sich eine der Zahlenfolgen unendlich oft gesetzt.

Die noch offene Kontinuumfrage einerseits und die erhebliche Unbestimmtheit in der Konstruktion solcher größter geordneter Mengen andererseits beschränken unsere Kenntnisse auf dieses Minimum.

Wir haben unser Beispiel absichtlich so formuliert, daß die Verbindung mit der Theorie der geordneten Mengen gewahrt blieb und eine allgemeine Untersuchung final geordneter Komplexmengen nach demselben Verfahren leicht durchgeführt werden kann. Andererseits, wenn es sich um reelle Komplexe (Komplexe, deren Elemente reelle Zahlen sind) handelt, so kann man natürlich spezielle Eigenschaften der reellen Zahlen, z. B. ihre Verknüpfungen durch die vier Spezies, ins Spiel ziehen und damit den Boden der allgemeinen Mengenlehre verlassen. Dies tun wir schon, wenn wir an der Definition der finalen Rangordnung die kleine Umformung anbringen, daß wir $a \gtreqless b$ definieren, wenn die Differenzen $a_n - b_n$ schließlich $\gtreqless 0$ werden; denn von einer Differenz kann bei Elementen einer beliebigen geordneten Menge nicht gesprochen werden. Wir tun es ebenso, wenn wir, unter Beschränkung auf Folgen positiver Zahlen, $a \gtreqless b$ definieren, falls der Quotient a_n / b_n schließlich $\gtreqless 1$ wird. Auf diesem Standpunkt lassen sich natürlich noch viele andere Ordnungen definieren, von denen wir eine in der Literatur besonders häufige hervorheben und durch eingeklammerte Ordnungszeichen andeuten wollen: man sagt, unter Beschränkung auf Folgen positiver Zahlen, daß

$$a(<)b, \quad a(=)b, \quad a(>)b,$$

wenn

$$\lim \frac{a_n}{b_n} = 0, \quad \lim \frac{a_n}{b_n} \neq 0, \quad \lim \frac{b_n}{a_n} = 0,$$

während man $a(\|)b$ definiert, wenn weder $\lim \frac{a_n}{b_n}$ noch $\lim \frac{b_n}{a_n}$ existiert. Sei es, daß man diese „infinitäre" Ordnung direkt untersucht, sei es, daß man sie auf unsere finale Ordnung zurückführt[1], erkennt man leicht, daß sich nichts Wesentliches ändert und insbesondere der Satz I mit seinen Konsequenzen bestehen bleibt. Der Leser übersieht jetzt, daß alle Probleme, die sich auf Graduierung des schließlichen Verlaufs von Zahlenfolgen beziehen, sei es ihres Null- oder Unendlichwerdens, ihrer Konvergenz oder Divergenz u. dgl., in der Hauptsache sich dem Schema unseres obigen Beispiels fügen, und daß die eigentümlichen Schwierigkeiten dieser Probleme auf der eventuellen Unvergleichbarkeit und unserer geringen Bekanntschaft mit der Struktur der größten geordneten Mengen beruhen. Genau dasselbe gilt, wenn man die Zahlenfolgen durch Funktionen

---

[1] Z. B. bedeutet $a(<)b$, daß für jede natürliche Zahl m die Zahlenfolge $ma = (ma_1, ma_2, \ldots)$ final kleiner als die Zahlenfolge b ist.

einer reellen Variablen ersetzt, und die Untersuchungen über Wachstum von Funktionen, Konvergenz und Divergenz uneigentlicher bestimmter Integrale usw. zeigen ebenfalls denselben Typus wie unsere obige Betrachtung über finale Ordnung von Zahlenfolgen.

§ 11. Komplexe reeller Zahlen.

Wenn wir Komplexe $a = (\ldots, a_i, \ldots)$ mit dem Argument J betrachten und voraussetzen, daß die Elemente a_i reelle Zahlen sind, so lassen sich, wie wir in § 10 hervorhoben, die Verknüpfungen reeller Zahlen auch auf diese Zahlenkomplexe übertragen, wenigstens zum Teil. So liegt es nahe, die Summe resp. Differenz zweier Komplexe als Komplex der Summen resp. Differenzen zu definieren:

$$(1) \qquad \begin{cases} a + b = (\ldots, a_i + b_i, \ldots, a_k + b_k, \ldots) \\ a - b = (\ldots, a_i - b_i, \ldots, a_k - b_k, \ldots). \end{cases}$$

Das Analoge könnte auch noch für das Produkt geschehen, also

$$ab = (\ldots, a_i b_i, \ldots, a_k b_k, \ldots),$$

für den Quotienten

$$\frac{a}{b} = \left(\ldots, \frac{a_i}{b_i}, \ldots, \frac{a_k}{b_k}, \ldots\right)$$

indessen nur dann, wenn kein Element b_i Null ist. Die Division ist damit also nur sehr unvollkommen definiert, denn das Ideal wäre doch, nur einen einzigen Komplex, den Nullkomplex $(\ldots, 0, \ldots, 0, \ldots)$, der aus lauter Nullen besteht, als Divisor ausschließen zu müssen, während hier schon alle Komplexe, die auch nur eine Null enthalten, auszuschließen sind. Aus diesem Grunde und noch anderen spielt die soeben definierte Multiplikation und Division keine hervorragende Rolle, und man hat sich nach geeigneteren Definitionen umgesehen. Wir erinnern den Leser hier nur an die Multiplikation und Division von Zahlenpaaren $a = (a_1, a_2)$:

$$ab = (a_1 b_1 - a_2 b_2, \quad a_1 b_2 + a_2 b_1),$$

$$\frac{a}{b} = \left(\frac{a_1 b_1 + a_2 b_2}{b_1^2 + b_2^2}, \quad \frac{a_2 b_1 - a_1 b_2}{b_1^2 + b_2^2}\right) \quad \text{für } b \neq (0, 0),$$

die in der Tat nur die Division durch das Nullpaar $(0, 0)$ ausschließt und ersichtlich nichts anderes ist, als die Multiplikation und Division der gewöhnlichen komplexen Zahlen $a_1 + i a_2$: in dieser Weise, als Paare reeller Zahlen, werden ja die komplexen Zahlen nach dem Vorgang von R. Hamilton heute in allen Lehrbüchern und Vorlesungen definiert. Wir erinnern weiter daran, daß es noch ein System von reellen Zahlenkomplexen mit vier Elementen, die Hamiltonschen Quaternionen $a = (a_1, a_2, a_3, a_4)$ gibt, zwischen denen sich ebenfalls eine allerdings nichtkommutative Multiplikation mit

entsprechenden Divisionen definieren läßt, mit alleiniger Ausschließung der Division durch $(0, 0, 0, 0)$. Nach einem Satze von G. Frobenius sind die reellen Zahlen, die gewöhnlichen komplexen Zahlen und die Quaternionen die einzigen Systeme von Zahlenkomplexen mit endlichem Argument, die nur die Division durch den Nullkomplex ausschließen.[1] Sieht man von dieser Forderung wieder ab, so gibt es schon im Falle endlichen Argumentes sehr viele und verschiedene Arten, Multiplikation und Division (bei nichtkommutativer Multiplikation zwei Divisionen) zu definieren. Die Theorie der Systeme komplexer Zahlen, in der diese Fragen behandelt werden, liegt uns hier natürlich fern, und darum werden wir auf unserm Standpunkt darauf verzichten müssen, auch die Multiplikation und Division von Zahlenkomplexen in den Kreis unserer allgemeinen Betrachtung zu ziehen.

Statt dessen heben wir ein anderes Moment hervor, das der Ordnung, das wieder in der eben erwähnten Theorie von geringerer Bedeutung ist. Wir kehren dabei zur lexikographischen Ordnung zurück. Wenn wir dann eine geordnete Menge von Zahlenkomplexen mit der durch (1) definierten Addition und Subtraktion betrachten, den Nullkomplex $(\ldots, 0, \ldots, 0, \ldots)$ einfach mit 0 bezeichnen und einen Komplex $a > 0$ positiv, einen Komplex $a < 0$ negativ nennen, so stellt sich schon im Fall von Zahlenpaaren eine Erscheinung ein, die im Gebiete der reellen Zahlen nicht auftritt: es kann nämlich ein positives Element a nicht nur größer sein als ein positives Element b, sondern auch größer als jedes Vielfache von b, d. h. größer als $b + b$, $b + b + b$, ... Das ist z. B. der Fall mit $a = (1, 0)$ und

$$b = (0, 1), \; b + b = (0, 2), \; b + b + b = (0, 3), \; \ldots.$$

Ein solches System nennt man nichtarchimedisch; indem man üblicherweise den Tatbestand, daß ein positives Element nach genügend häufiger Vervielfachung jedes andere positive Element übertrifft, als das Axiom des Archimedes bezeichnet. Man kann das obige Verhalten von a und b auch dadurch charakterisieren, daß man a unendlich groß gegen b, b unendlich klein gegen a nennt.

Man sieht unmittelbar, daß im Falle eines beliebigen Argumentes J von den positiven Komplexen a, b stets a unendlich groß gegen b wird, sobald der kritische Index (die erste Differenzstelle) zwischen a und dem Nullkomplex dem kritischen Index zwischen b und dem Nullkomplex vorangeht. Denn sind i, k diese kritischen

[1] Natürlich sind hierbei gewisse Bedingungen hinzuzufügen, denen die Verknüpfungen unterworfen werden.

Indices und $i < k$, so ist $a_h = 0$ für $h < i$, $a_i > 0$, während die Komplexe $b, b+b, b+b+b, \ldots$ an allen Stellen $\leqq i$ Nullen tragen, also lexikographisch $< a$ sind.

Um diese Frage noch etwas eingehender zu behandeln, sehen wir zunächst von unseren Komplexen ab und definieren ein Größensystem als eine geordnete Menge, zwischen deren Elementen (Größen) eine folgenden Bedingungen genügende Addition definiert ist:

Je zwei Elemente α, β des Systems bestimmen eindeutig ein drittes Element $\alpha+\beta$ des Systems. Diese Addition ist kommutativ und assoziativ:

$$\beta+\alpha=\alpha+\beta, \; (\alpha+\beta)+\gamma=\alpha+(\beta+\gamma).$$

Sie ist eindeutig umkehrbar, also zwei Elemente α, β des Systems bestimmen eindeutig ein drittes Element $\beta-\alpha$ des Systems, für welches

$$\alpha+(\beta-\alpha)=\beta.$$

Endlich soll aus $\alpha<\beta$ stets $\gamma+\alpha<\gamma+\beta$ folgen. Eine solche Addition heiße kurz eine normale Addition.

Es ist stets $\alpha-\alpha=\beta-\beta$; dieses Element des Systems wird mit 0 bezeichnet, so daß, für jedes Element α, $\alpha+0=\alpha$ ist.

Das Element $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ wird auch mit $\alpha+\beta+\gamma$ bezeichnet, und analog für jede endliche Zahl von Summanden. Man schreibt

$$1\alpha=\alpha, \; 2\alpha=\alpha+\alpha, \; 3\alpha=\alpha+\alpha+\alpha, \; \ldots;$$

ferner wird[1] $0\alpha=0$, $-\alpha=0-\alpha$ und für jede natürliche Zahl n

$$(-n)\alpha=-(n\alpha)=0-(n\alpha)$$

gesetzt, womit $m\alpha$ für jede ganze Zahl m definiert ist. Jede ganzzahlige lineare Verbindung $m_1\alpha_1+m_2\alpha_2+\ldots+m_n\alpha_n$ von Elementen des Systems ist wieder ein Element des Systems.

Aus $\alpha<\beta$ folgt

$$2\alpha=\alpha+\alpha<\alpha+\beta<\beta+\beta=2\beta,$$

und ebenso $n\alpha<n\beta$ für jede natürliche Zahl n. Aus $n\alpha=n\beta$ folgt also umgekehrt $\alpha=\beta$; demnach gibt es zu einem Element γ des Systems höchstens ein Element α derart, daß $n\alpha=\gamma$. Wenn ein solches im System vorhanden ist, so bezeichnen wir es mit $\alpha=\frac{1}{n}\gamma$. Wenn nicht, so läßt sich das Größensystem stets zu einem solchen erweitern, in dem auch jeder Bruchteil eines jeden Elements

[1] Hier ist 0 links die Zahl 0, rechts das Element 0 des Größensystems; diese Zweideutigkeit ist unschädlich.

vorkommt. Denn man bilde aus den Elementen des Systems und den natürlichen Zahlen die Paare (α, p) und definiere

$$(\alpha, p) \gtreqqless (\beta, q) \quad \text{für} \quad q\alpha \gtreqqless p\beta,$$

ferner

$$(\alpha, p) + (\beta, q) = (q\alpha + p\beta, pq),$$

so bilden, wie leicht zu sehen, diese Paare wieder ein Größensystem, in dem zu jedem Element (α, p) auch der n^{te} Teil (α, pn) vorhanden ist; den Elementen α des alten Systems entsprechen umkehrbar eindeutig, unter Erhaltung von Addition und Ordnung, die Paare $(\alpha, 1)$ des neuen. Wir wollen zur Vereinfachung gleich das ursprüngliche System als ein solches mit Teilbarkeit der Elemente annehmen. Dann ist für jedes Element α auch $\frac{1}{n}\alpha$ definiert, ebenso für jede rationale Zahl $r = \frac{m}{n}$ mit positivem Nenner

$$r\alpha = m\beta, \; \beta = \frac{1}{n}\alpha.$$

Jede rationalzahlige lineare Verbindung $r_1\alpha_1 + r_2\alpha_2 + \dots + r_n\alpha_n$ von Elementen des Systems ist wieder ein Element des Systems.

Sind α, β zwei positive Elemente $(\alpha > 0, \beta > 0)$, so sind drei Fälle möglich:

Entweder ist, für jede natürliche Zahl n, $n\alpha < \beta$. Dann heißt α unendlich klein gegen β, wofür wir $\alpha (<) \beta$ schreiben; insbesondere ist dann auch $\alpha < \beta$.

Oder für jede natürliche Zahl n ist $\alpha > n\beta$. Dann heißt α unendlich groß gegen β, geschrieben $\alpha (>) \beta$. Dieser Fall, wo $\alpha > \beta$, schließt den vorigen aus.

Oder drittens, es tritt keiner der beiden ersten Fälle ein; α heiße endlich gegen β, geschrieben $\alpha (=) \beta$. Es gibt also dann natürliche Zahlen p, q derart, daß $p\alpha \geqq \beta$ und $\alpha \leqq q\beta$, von denen die eine gleich 1 angenommen werden kann.

Wenn die beiden ersten Fälle niemals eintreten, so ist das System archimedisch, andernfalls nichtarchimedisch.

Für Elemente α, β, die positiv oder negativ (aber nicht Null) sind, wollen wir $\alpha (\lesseqqgtr) \beta$ definieren, je nachdem für ihre Beträge $|\alpha| (\lesseqqgtr) |\beta|$ ist; unter dem Betrag $|\alpha|$ verstehen wir das positive von den beiden Elementen $\pm\alpha$. Das Element 0 können wir, wenn wir wollen, als unendlich klein gegen jedes andere Element definieren oder auch von der Vergleichung ausschließen.

Die formalen Gesetze dieser Zeichen sind dieselben wie gewöhnlich. Aus den transitiven Gesetzen schließt man: alle gegen ein Element $\alpha \neq 0$ endlichen Elemente sind auch gegeneinander endlich; wird ihre Menge als eine Größenklasse $K(\alpha)$ bezeichnet,

so haben zwei verschiedene Größenklassen kein Element gemein; ist α unendlich klein gegen β, so ist jedes Element der Klasse $K(\alpha)$ unendlich klein gegen jedes Element der Klasse $K(\beta)$. Sagt man in diesem Fall, daß die Klasse $K(\alpha)$ der Klasse $K(\beta)$ vorangehe oder daß $K(\alpha)$ die niedere, $K(\beta)$ die höhere Klasse sei, so bilden die Klassen eine geordnete Menge, die im Fall eines nichtarchimedischen Systems aus mindestens zwei Elementen besteht, während ein archimedisches nur eine Klasse liefert.[1] Die Summe aller Klassen ist die Menge der von 0 verschiedenen Größen des Systems.

Prüfen wir auf Grund dieser Festsetzungen unsere Komplexmengen mit dem Argument J, so zeigt sich unmittelbar, daß wir, und zwar noch in verschiedener Weise, nichtarchimedische Größensysteme erhalten, deren Klassenmenge mit dem Exponenten $E = J^*$ ähnlich ist. Die Ordnung soll, wie gesagt, lexikographisch und Addition nebst Subtraktion durch (1) definiert sein. Das Größensystem, das wir aus Komplexen bilden wollen, muß jedenfalls, wenn a eins seiner Elemente ist, den Komplex $a - a$, d. h. den Nullkomplex $0 = (\ldots, 0, \ldots)$ enthalten; um dann eine geordnete Menge zu konstruieren, verfahren wir nach den Prinzipien in § 3 und bilden zunächst die Menge A aller mit 0 kongruenten Komplexe a, derer also, für welche die Menge der Indices, wo $a_i \neq 0$, eine wohlgeordnete Teilmenge von J ist. Das ist also nichts anderes als die Maximalpotenz

$$A = (M, 0)^{J^*}$$

mit der Menge M der reellen Zahlen als Basis und der Zahl 0 als Hauptelement. Da aus $a \equiv 0$, $b \equiv 0$ auch $a \pm b \equiv 0$ folgt, so ist im System A eine unseren Forderungen genügende Addition definiert, eine normale Addition also, da offenbar mit $a < b$ zugleich $c + a < c + b$ ist. Ist $f(a)$ der kritische Index zwischen a und dem Nullkomplex und etwa $a > b > 0$, so ist $f(a) \leqq f(b)$; wir wiesen bereits darauf hin, daß für $f(a) < f(b)$ gewiß b unendlich klein gegen a ist, wovon offenbar auch die Umkehrung gilt. Allgemein, für zwei von 0 verschiedene Komplexe, ist also $a \, (\lesseqqgtr) \, b$, je nachdem $f(a) \gtreqqless f(b)$; jedem Index i entspricht also eine Größenklasse K_i, nämlich die Menge aller Komplexe a, für die $f(a) = i$, und diese Klassen haben die umgekehrte Rangordnung der Indices in J oder dieselbe Rangordnung wie die Indices in $J^* = E$. Die Klassenmenge ist daher mit E ähnlich, und man sieht, daß man nichtarchimedische Größensysteme mit beliebig vorgeschriebenem Typus ihrer Klassenmenge bilden kann.

Die Maximalpotenz ist aber nicht das einzige System dieser

---

[1] Das Element 0 rechnen wir keiner Klasse zu.

Art. Auch Teilpotenzen vom Grade ξ leisten dasselbe, d. h. die Menge

$$A_\xi = (M, 0)_\xi^{J*}$$

aller Komplexe a, die für die reguläre Anfangszahl ω_ξ mit dem Nullkomplex kongruent sind, ist wieder ein Größensystem mit einer Klassenmenge $\simeq E$. Das beruht darauf, daß aus $a \equiv 0\,(\omega_\xi)$, $b \equiv 0\,(\omega_\xi)$ wieder $a \pm b \equiv 0\,(\omega_\xi)$ folgt.

Weiter kann auch die Basis noch verändert werden; es genügt, statt der Menge der reellen Zahlen die der rationalen Zahlen oder der Zahlen $r_1 + r_2\sqrt{2}$ (r_1, r_2 beliebige rationale Zahlen) u. dgl. zu wählen. Das allgemeinste für M zulässige System reeller Zahlen muß offenbar so beschaffen sein, daß in ihm Addition und Subtraktion und, falls wir Teilbarkeit der Komplexe verlangen, auch die Division durch natürliche Zahlen ausführbar ist. Ist X eine beliebige Menge reeller Zahlen und verstehen wir unter μX den Inbegriff der Zahlen, die als ganzzahlige lineare Verbindungen

$$m_1 x_1 + m_2 x_2 + \dots + m_n x_n$$

von endlich vielen Zahlen $x_1, x_2, \dots, x_n$ der Menge X darstellbar sind, unter ϱX den Inbegriff der Zahlen, die ebenso als rationalzahlige lineare Verbindungen

$$r_1 x_1 + r_2 x_2 + \dots + r_n x_n$$

darstellbar sind, so muß demnach $M = \mu M$, resp. $M = \varrho M$ sein. Dazu ist notwendig, daß M ein System der Form μX, resp. ϱX sei; dies ist aber auch hinreichend, da $\mu\mu X = \mu X$ und $\varrho\varrho X = \varrho X$ ist. Es ist also für M irgend ein System μX resp. ϱX zulässig, bei beliebiger[1] Menge X; und schließlich kann man auch die Komplexelemente a_i verschiedene Mengen A_i dieser Art durchlaufen lassen, also von Potenzen zu Produkten übergehen. Allgemein also ist zu sagen, daß jene Maximalpotenz $A = (M, 0)^E$ (mit der Menge der reellen Zahlen als Basis und der Zahl 0 als Hauptelement) und geeignete Teilmengen von ihr Größensysteme mit einer Klassenmenge $\simeq E$ liefern.

Es ist interessant und wirft ein neues Licht auf unseren Produkt- und Potenzbegriff, daß umgekehrt, wie H. Hahn bewiesen hat, die zuletzt genannten Größensysteme auch die allgemeinsten ihrer Art sind, oder präziser: man kann jedes Größensystem mit einer Klassenmenge $\simeq E$ auf eine Teilmenge der Maximalpotenz $(M, 0)^E$ umkehrbar eindeutig unter Erhaltung von Rangordnung und Addition abbilden. Wenn also

[1] Natürlich soll X weder die Nullmenge sein noch aus der einzigen Zahl 0 bestehen.

ein beliebiges Größensystem vorliegt, dessen Klassenmenge zu E ähnlich ist, so kann man seinen Größen $\alpha, \beta, \ldots$ Komplexe $a, b, \ldots$ mit dem Argument $J = E^*$ zuordnen, die sämtlich mit dem Nullkomplex kongruent sind, d. h. nur an einer wohlgeordneten Teilmenge von J nichtverschwindende Elemente aufweisen, und zwar so, daß mit $\alpha \lesseqgtr \beta$ zugleich $a \lesseqgtr b$ ist und daß der Größe $\alpha + \beta$ der Komplex $a + b$ entspricht.

Um dies zu beweisen, nehmen wir wieder das Größensystem Γ derart an, daß es unbeschränkte Teilung seiner Elemente gestattet (wenn der Satz für ein solches gilt, gilt er natürlich auch für jedes Teilsystem von Γ). Ist Δ eine beliebige Teilmenge von Γ, so bezeichnen wir die Menge der Größen, die sich als rationalzahlige lineare Verbindungen

$$r_1 \alpha_1 + r_2 \alpha_2 + \ldots + r_n \alpha_n$$

von endlich vielen Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ aus Δ darstellen lassen, mit $\varrho \Delta$, analog einer vorhin angewandten Bezeichnung. Ist ebenso D eine beliebige Menge von Komplexen, so bedeute ϱD die Menge der Komplexe

$$r_1 a_1 + r_2 a_2 + \ldots + r_n a_n,$$

die aus endlich vielen Komplexen $a_1, a_2, \ldots, a_n$ von D mit rationalen Koeffizienten zusammengesetzt sind.[1]

Wir nennen die Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ abhängig, wenn eine Relation

$$r_1 \alpha_1 + r_2 \alpha_2 + \ldots + r_n \alpha_n = 0$$

zwischen ihnen besteht, deren rationale Koeffizienten (die man hier auch ganzzahlig annehmen kann) nicht sämtlich verschwinden; andernfalls unabhängig. Eine Größenmenge $\Delta \leqq \Gamma$ heißt unabhängig, wenn je n verschiedene ihrer Größen unabhängig sind ($n = 1, 2, 3, \ldots$). Es ist evident, daß es größte unabhängige Teilmengen von Γ gibt; der Beweis dafür ist genau ebenso wie der für die Existenz größter geordneter Teilmengen einer teilweise geordneten Menge (§ 1) zu führen, nämlich: man wende das Zermelosche Wohlordnungsverfahren in der Weise an, daß als ausgezeichnetes Element von Γ', wenn möglich, ein Element von $\Gamma - \varrho \Gamma''$ ($\Gamma'' = \Gamma - \Gamma'$) gewählt wird. Hierdurch wird eine Wohlordnung

$$\Gamma = \{\alpha_0, \alpha_1, \ldots\}$$

definiert. Da Γ selbst kein unabhängiges System ist, so gibt es ein erstes Element α_λ, das als rationalzahlige lineare Verbindung

[1] Natürlich ist, für $a = (\ldots, a_i, \ldots, a_k, \ldots)$, unter ta der Komplex $(\ldots, ta_i, \ldots, ta_k, \ldots)$ zu verstehen, wobei t irgend eine reelle Zahl sein kann.

vorangehender Elemente (d. h. solcher mit kleineren Indices) darstellbar ist. Die Menge dieser vorangehenden Elemente

(2) $$\varDelta = \{\alpha_0, \alpha_1, \ldots, \alpha_\xi, \ldots\} \qquad (\xi < \lambda)$$

ist dann unabhängig; sie ist ferner eine größte unabhängige Teilmenge, denn α_λ ist das ausgezeichnete Element von $\Gamma - \varDelta$, sollte also als Element von $\Gamma - \varrho\varDelta$ gewählt werden, wenn möglich, d. h. wenn diese Menge von Null verschieden ist. Da α_λ aber dieser Menge nicht (sondern der Menge $\varrho\varDelta$) angehört, so muß sie 0 sein; demnach ist $\Gamma = \varrho\varDelta$, und jede Größe des Systems ist also durch Größen aus $\varDelta$ linear mit rationalen Koeffizienten ausdrückbar. Wir nennen eine solche größte unabhängige Teilmenge $\varDelta$ eine Basis von Γ. Es ist klar, daß die Darstellung einer beliebigen Größe α durch die Größen der Basis, abgesehen von Hinzufügung oder Weglassung von Gliedern mit verschwindenden Koeffizienten, nur auf eine Weise möglich ist; d. h. schreibt man

(3) $$\alpha = \sum r_\xi \alpha_\xi,$$

wobei aber nur eine endliche Anzahl von Koeffizienten r_ξ von 0 verschieden ist, so sind die r_ξ durch α eindeutig bestimmt, da sich andernfalls durch Subtraktion eine Abhängigkeit zwischen den Größen der Basis ergeben würde.

Unser Problem, die Größenmenge Γ unter Erhaltung von Addition und Ordnung auf eine Komplexmenge C abzubilden, vereinfacht sich durch Einführung einer Basis $\varDelta$ wesentlich. Wir brauchen nur den Größen der Basis Komplexe zuzuordnen, deren Menge

(4) $$D = \{a_0, a_1, \ldots, a_\xi, \ldots\} \qquad (\xi < \lambda)$$

unabhängig ist, und dann der Größe (3) den Komplex

(5) $$a = \sum_\xi r_\xi a_\xi$$

entsprechen zu lassen. Für $\alpha \neq \beta$ ist dann auch $a \neq b$, und der Größe $\alpha + \beta$ entspricht der Komplex $a + b$. Wenn wir ferner jeden Komplex $a_\xi = 0$ wählen, so ist auch $a = 0$. Damit wäre also Γ eindeutig umkehrbar und mit Erhaltung der Addition auf eine Teilmenge C jener Maximalpotenz $A = (M, 0)^E$ abgebildet, falls $J = E^*$ das Argument der Komplexe ist. Was die Erhaltung der Ordnung, die Ähnlichkeit der Abbildung betrifft, so ist zu bemerken: zunächst genügt es, daß jeder positiven Größe ein positiver Komplex, jeder negativen Größe ein negativer Komplex entspricht, denn dann folgt aus $\alpha - \beta \lesseqgtr 0$ auch $a - b \lesseqgtr 0$, aus $\alpha \lesseqgtr \beta$ also $a \lesseqgtr b$. Ist ferner

in der Darstellung (3) η die größte Ordnungszahl, der ein nichtverschwindender Koeffizient r_η entspricht, so ist ($\xi < \eta$)

$$\alpha = \sum_\xi r_\xi \alpha_\xi + r_\eta \alpha_\eta = r_\eta(\alpha_\eta - \beta),$$

wobei

$$\beta = -\sum_\xi \frac{r_\xi}{r_\eta} \alpha_\xi$$

eine aus Größen der Menge

(6) $$\Delta_\eta = \{\alpha_0, \alpha_1, \ldots, \alpha_\xi, \ldots\} \qquad (\xi < \eta)$$

gebildete rationalzahlige lineare Verbindung, also eine Größe aus $\varrho \Delta_\eta$ ist. Das Vorzeichen von α ist, je nachdem r_η positiv oder negativ ist, dasselbe oder das entgegengesetzte wie das von $\alpha_\eta - \beta$. Da für die entsprechenden Komplexe dasselbe gilt, so folgt: um die Abbildung ähnlich zu machen, genügt es, daß $a_\eta - b$ dasselbe Zeichen wie $\alpha_\eta - \beta$ habe, wo b der der Größe β entsprechende Komplex aus ϱD_η und

(7) $$D_\eta = \{a_0, a_1, \ldots, a_\xi, \ldots\} \qquad (\xi < \eta)$$

ist, d. h. es genügt, jeder Größe α_η der Basis einen Komplex a_η zuzuordnen, der zu allen Komplexen b der Menge ϱD_η dieselbe Ordnung hat wie α_η zu den Größen β der Menge $\varrho \Delta_\eta$. Es ist also induktiv zu zeigen, daß eine solche Wahl für η möglich ist, wenn sie für jedes $\xi < \eta$ bereits geglückt, d. h. die Menge ϱD_η auf $\varrho \Delta_\eta$ ähnlich abgebildet ist. Die oben noch aufgestellte Bedingung der Unabhängigkeit der Menge D ist bei diesem Verfahren von selber gesichert und braucht also nicht weiter beachtet zu werden.

Dagegen spezialisieren wir die Abbildung noch durch die folgende Vorschrift. Die Klassenmenge des Größensystems war mit J^* ähnlich angenommen; es entspricht also jedem Index i eine Größenklasse K_i, wobei aber die Klassen die umgekehrte Ordnung wie die Indices in J haben. Den Klassenindex einer Größe $\alpha \neq 0$ des Systems bezeichnen wir mit $f(\alpha)$, wobei also $i = f(\alpha)$ besagt, daß α zur Klasse K_i gehört. Es ist $f(\alpha) \lesseqgtr f(\beta)$, je nachdem $\alpha \gtreqless \beta$. Andererseits soll, für einen Komplex $a \neq 0$, $f(a)$ wie bisher den kritischen Index von a mit dem Komplex 0 bedeuten; $f(a) = i$ besagt also, daß a_i das erste nichtverschwindende Element von a ist. Die Abbildung soll dann die Forderung $f(\alpha) = f(a)$ erfüllen, einer Größe der Klasse K_i soll ein Komplex mit dem kritischen Index i entsprechen; wir werden sofort zeigen, daß man diese Bedingung erfüllen kann.

Bezüglich der Wahl der Basis sind noch mancherlei spezielle Annahmen gestattet. So können wir alle Basisgrößen positiv annehmen; ferner ist, wenn β_η irgend eine Größe aus $\varrho \Delta_\eta$ bedeutet, leicht einzusehen, daß auch die Größen $\alpha_\eta - \beta_\eta$ eine Basis bilden. Endlich kann man erreichen, daß Δ eine vorgegebene unabhängige

Größenmenge Δ' als Teilmenge und sogar als Abschnitt enthält (indem man, wie in § 1, bei der Wohlordnung vorschreibt, daß das ausgezeichnete Element von Γ', wenn möglich, zu Δ' gehören soll).

Wählt man aus jeder Klasse K_i ein positives Element ε_i, so bilden diese offenbar ein unabhängiges[1] Größensystem Δ'; wir lassen Δ mit dem Abschnitt Δ' beginnen. Wir ordnen ferner der Größe ε_i den Komplex e_i zu, der an der Stelle i die Zahl Eins und sonst lauter Nullen trägt; die Menge dieser „Einheitskomplexe" sei D'. Es ist evident, daß die Menge $\varrho\Delta'$ auf die Menge $\varrho D'$ ähnlich abgebildet wird; denn die Größe

$$r_i\varepsilon_i + r_k\varepsilon_k + \dots + r_l\varepsilon_l,$$

wo $i < k < \dots < l$ in J vorausgesetzt sei und die Koeffizienten nicht verschwinden, gehört zur Klasse K_i und hat das Vorzeichen von r_i, dasselbe Vorzeichen hat aber auch der zugeordnete Komplex

$$r_ie_i + r_ke_k + \dots + r_le_l,$$

der an den Stellen $i, k, \dots, l$ die Elemente $r_i, r_k, \dots, r_l$ und sonst lauter Nullen trägt. Hiermit ist auch erreicht, daß bei der ähnlichen Abbildung (falls sie überhaupt möglich ist oder soweit sie möglich ist) die oben gestellte Forderung $f(\alpha) = f(a)$ realisiert wird. Denn ist α, das wir als positiv voraussetzen dürfen, ein Element der Klasse K_i, also $\alpha (=) \varepsilon_i$, so gibt es zwei positive rationale Zahlen r, s derart, daß $r\varepsilon_i < \alpha < s\varepsilon_i$; da dann auch $re_i < a < se_i$ ist, so ist der kritische Index $f(a)$ gleichzeitig $\leqq f(re_i) = i$ und $\geqq f(se_i) = i$, also gleich i.

Nach diesem Anfang der Abbildung gestaltet sich nun der Induktionsbeweis für die Möglichkeit ihrer Fortsetzung[2] folgendermaßen. Es sei bereits $\varrho\Delta_\eta$ auf ϱD_η ähnlich abgebildet, wobei also jeder Größe β der ersten Menge ein Komplex $b = 0$ der zweiten entspricht und $f(\beta) = f(b)$ ist. Wir zeigen, daß man dann auch der Basisgröße α_η einen Komplex $a_\eta = 0$ zuordnen kann, der zu den Komplexen b dieselbe Ordnung hat wie α_η zu den Größen β. Schreiben wir kurz α für α_η und a für a_η. Es sind zwei Fälle zu unterscheiden:

I. Unter den Klassen aller Größen $\alpha - \beta$ (wobei β seine Menge $\varrho\Delta_\eta$ durchläuft) sei eine niedrigste K_i.

Es ist also stets $\alpha - \beta (\geqq) \varepsilon_i$ und, für ein bestimmtes β_0,

---

[1] Gehören α und β zu einer Klasse, so gehört $\alpha + \beta$ zur selben oder zu einer niederen Klasse; ist hingegen β unendlich klein gegen α, so gehört $\alpha + \beta$ zur selben Klasse wie α. Ist $\alpha (>) \beta (>) \gamma (>) \dots$, so ist $\alpha + \beta + \gamma + \dots (=) \alpha$. Größen verschiedener Klassen sind danach stets unabhängig.

[2] wobei also η bereits so groß angenommen wird, daß $\Delta' \leqq \Delta_\eta$.

$\alpha - \beta_0 (=) \varepsilon_i$. Indem wir, nach der obigen Bemerkung, α durch $\alpha - \beta_0$ ersetzen, können wir erreichen, daß α selbst von der niedrigsten Klasse K_i, also

$$\alpha - \beta (\geqq) \varepsilon_i (=) \alpha$$

wird; überdies setzen wir α positiv voraus. Unter r verstehen wir eine willkürliche positive rationale Zahl.

Da infolge des Anfangs unserer Abbildung $r\varepsilon_i$ ein β, also niemals $\alpha = r\varepsilon_i$ ist, so ist $\alpha \lesseqgtr r\varepsilon_i$. Es kann weder stets $\alpha < r\varepsilon_i$, d. h. $\alpha (<) \varepsilon_i$, noch stets $\alpha > r\varepsilon_i$ sein, und wir erhalten also eine Einteilung der r in zwei Klassen derart, daß

$$u\varepsilon_i < \alpha < v\varepsilon_i$$

und $u < v$. Es kann kein größtes u geben, denn für dieses u und jedes r wäre

$$u\varepsilon_i < \alpha < (u + r)\varepsilon_i,$$
$$0 < \alpha - u\varepsilon_i < r\varepsilon_i,$$

also $\alpha - u\varepsilon_i (<) \varepsilon_i$, während doch stets $\alpha - \beta (\geqq) \varepsilon_i$ sein soll. Ebenso gibt es keine kleinste Zahl v. Die genannte Einteilung bestimmt also eine irrationale Zahl q derart, daß $u < q < v$, und nun ordnen wir der Größe α den Komplex

$$a = qe_i$$

zu, der an der Stelle i die Zahl q und sonst lauter Nullen hat, also mit dem Nullkomplex kongruent ist.

Daß a zu den Komplexen b die richtige Ordnung hat, ist leicht einzusehen. Sei etwa β eine Größe $< \alpha$. Für die oben genannten Zahlen u kann dann nicht stets $u\varepsilon_i < \beta$ sein, denn aus

$$u\varepsilon_i < \beta < \alpha < v\varepsilon_i$$

würde

$$0 < \alpha - \beta < (v - u)\varepsilon_i,$$

also $\alpha - \beta (<) \varepsilon_i$ folgen. Es gibt also sicher eine Zahl u derart, daß $u\varepsilon_i \geqq \beta$. Da die Größen β, nach Annahme, durch die zugehörigen Komplexe b ähnlich abgebildet werden, so ist $ue_i \geqq b$, außerdem $qe_i > ue_i$, also $qe_i > b$ oder $a > b$. Ebenso folgt aus $\alpha < \beta$ zugleich $a < b$.

II. Unter den Klassen der Größen $\alpha - \beta$ sei keine niedrigste.

Oder unter den Klassenindices $f(\alpha - \beta)$ ist kein letzter. Wir bezeichnen die Menge dieser Klassenindices etwa mit K, mit

$$H = \mathfrak{S}_k^K J^k$$

die Summe der zugehörigen Anfangsstrecken von J oder die Menge der Indices h, zu denen es einen späteren Klassenindex $k = f(\alpha - \beta) > h$

gibt. H ist mit K konfinal, ohne letztes Element und ist ein Anfangsstück $\leqq J$. Wir haben hier ein ähnliches Raisonnement wie in § 5 für das gleichbezeichnete Anfangsstück H anzustellen.

Ist nämlich h ein Element von H, β eine Größe, für die $f(\alpha-\beta) > h$ und b der zugehörige Komplex, so behaupten wir, daß b an der Stelle h ein nur von dieser, nicht von b abhängiges festes Element a_h trägt, oder daß für zwei verschiedene solche Größen β und β' die zugehörigen Komplexe b und b' an der Stelle h übereinstimmen. In der Tat seien $f(\alpha-\beta)$ und $f(\alpha-\beta')$ beide $> h$. Da die Summe oder Differenz zweier Größen nie von höherer Klasse als beide Größen sein kann, so ist für die Größe $\alpha-\beta-(\alpha-\beta') = \beta'-\beta$ mindestens eine der beiden Ungleichungen

$$f(\beta'-\beta) \geqq f(\alpha-\beta) \quad \text{oder} \quad \geqq f(\alpha-\beta')$$

erfüllt, also sicher $f(\beta'-\beta) > h$. Für die zugehörigen Komplexe ist, da $\beta'-\beta$ zur Menge $\varrho\Delta_\eta$ gehört und also bereits sein richtiges Bild $b'-b$ erhalten hat, $f(b'-b) = f(\beta'-\beta) > h$, also wirklich $b_h' = b_h$. Wir bezeichnen diese nur von h abhängige reelle Zahl mit a_h.

Diese Zahlen a_h bilden einen Anfangskomplex $a_H = (\ldots, a_h, \ldots)$, der offenbar mit dem entsprechenden Anfange des Nullkomplexes kongruent ist. Denn ist $G = \{\ldots, h\}$ ein Anfangsintervall von H, so gibt es eine Größe β mit $f(\alpha-\beta) > h$, und für den zugehörigen Komplex b ist demnach

$$b_i = a_i \quad \text{für} \quad i \leqq h,$$

also $a_G = b_G$. Da nun $b \equiv 0$ vorausgesetzt war, so ist die Menge der Stellen in G, wo $a_i \neq 0$, wohlgeordnet; danach ist auch die Menge der Stellen in H, wo $a_i \neq 0$, wohlgeordnet, da jeder ihrer Abschnitte wohlgeordnet ist. Ergänzen wir, indem wir für die Stellen von $J-H$ lauter Nullen hinzufügen, a_H zu einem vollen Komplex a, so ist dieser mit 0 kongruent; ihn ordnen wir der Größe α als Bild zu.

Daß a zu den Komplexen b die richtige Ordnung hat, ergibt sich folgendermaßen. Da unter den Klassenindices $f(\alpha-\beta)$ kein letzter ist, so gehören sie zu dem Anfangsstück H. Es sei β irgend eine Größe (aus $\varrho\Delta_\eta$) und $f(\alpha-\beta) = h$; dann gibt es sicher auch eine Größe β' mit $f(\alpha-\beta') > h$. Hiernach ist $\alpha-\beta'$ unendlich klein gegen $\alpha-\beta$, und die Differenz

$$(\alpha-\beta)-(\alpha-\beta') = \beta'-\beta$$

gehört demzufolge derselben Klasse wie $\alpha-\beta$ an und hat dasselbe Vorzeichen wie diese Größe; es ist $f(\beta'-\beta) = h$ und $\beta' \gtreqless \beta$ für $\alpha \gtreqless \beta$. Die zu β, β' gehörigen Komplexe seien b, b'. Wegen $f(\alpha-\beta') > h$ ist, nach der Bestimmung der Zahlen a_i,

$$a_i = b_i' \quad \text{für} \quad i \leqq h.$$

Wegen $f(b'-b)=f(\beta'-\beta)=h$ ist

$$b_i'=b_i \text{ für } i<h, \quad b_h' \neq b_h,$$

also

$$a_i=b_i \text{ für } i<h, \quad a_h \neq b_h,$$

und zwar ist das Vorzeichen von $a_h-b_h=b_h'-b_h$ dasselbe wie das von $\beta'-\beta$ (weil die Abbildung der β auf die b ähnlich angenommen war) und also dasselbe wie das von $\alpha-\beta$. Demnach ist lexikographisch $a \gtreqless b$, je nachdem $\alpha \gtreqless \beta$.

Damit ist der Beweis des Hahnschen Satzes vollendet.

Zum Schluß kommen wir noch einmal auf die Multiplikation in Größensystemen zurück, auf deren allgemeine Behandlung wir ja verzichten wollten; wir möchten nur auf eine spezielle, der Multiplikation von Potenzreihen nachgebildete Produktdefinition hinweisen, die alle gewöhnlichen Multiplikationsgesetze befolgt und auch, bis auf die Division durch Null, eine eindeutige Umkehrung zuläßt. Legen wir als Größensystem die Maximalpotenz $(M, 0)^E$, d. h. die Menge aller aus reellen Zahlen[1] gebildeten, mit 0 kongruenten Komplexe vom Argument $J=E^*$ zugrunde und nehmen jetzt an, daß das Argument selber ein Größensystem mit normaler Addition ist. Wir definieren dann das Produkt $c=ab$ zweier Komplexe als denjenigen Komplex, der an der Stelle l die Zahl

$$(8) \qquad c_l=\Sigma a_i b_k \qquad (i+k=l)$$

trägt, wo also die Summe über alle Paare von Indices zu erstrecken ist, deren Summe l ist.

Um die Zulässigkeit dieser Definition zu erweisen, haben wir zunächst zu zeigen, das jede dieser Summen in Wahrheit nur endlich viele nichtverschwindende Glieder hat. Zunächst braucht i nur die wohlgeordnete Menge der Indices zu durchlaufen, wo $a_i \neq 0$, ebenso k nur die wohlgeordnete Menge der Indices, wo $b_k \neq 0$. Wegen der Eigenschaften der normalen Addition folgt aber aus

$$i+k=i'+k', \; i<i'$$

umgekehrt $k>k'$; die Menge der Indices k, die nichtverschwindende Glieder der Summe liefern, ist also der Menge der entsprechenden Indices i umgekehrt ähnlich, also gleichzeitig wohlgeordnet und invers wohlgeordnet, d. h. endlich. Damit ist für jedes l die reelle Zahl c_l definiert.

Ferner ist zu zeigen, daß mit a und b auch $c=0$ ist oder daß die Menge der Stellen l, wo $c_l \neq 0$, wohlgeordnet ist. Jedem solchen l

[1] Man könnte sich wieder auf die Menge der rationalen Zahlen oder sonstige geeignete Zahlenmengen beschränken.

entspricht mindestens ein Paar Indices i und k, für die $i+k=l$ und a_i, b_k von 0 verschieden sind. Gäbe es nun eine absteigende Folge solcher l, also eine Folge

$$i_1+k_1>i_2+k_2>i_3+k_3>\ldots,$$

wobei die i_n und k_n je einer wohlgeordneten Teilmenge von J angehören, so haben die Indices i_n ein Minimum i_p, ferner die Indices i_{p+n} ein Minimum i_q, die Indices i_{q+n} ein Minimum i_r usw., es gibt also eine Folge wachsender natürlicher Zahlen derart, daß

$$i_p+k_p>i_q+k_q>i_r+k_r>\ldots, i_p\leqq i_q\leqq i_r\leqq\ldots.$$

Daraus würde aber $k_p>k_q>k_r>\ldots$ folgen im Widerspruch dazu, daß auch die k_n einer wohlgeordneten Menge angehören.

Wir überlassen dem Leser den Nachweis, daß diese Multiplikation kommutativ, assoziativ und gegenüber der Addition distributiv ist.

Ein Produkt, das den Nullkomplex als Faktor enthält, ist selbst der Nullkomplex. Umgekehrt, wenn a und b von 0 verschieden sind, so ist auch $c=ab$ von 0 verschieden. Führt man nämlich die kritischen Indices (gegenüber dem Nullkomplex) $f(a)=i_0$ und $f(b)=k_0$ ein und setzt $l_0=i_0+k_0$, so ist für $i+k=l<l_0$ mindestens eine der Ungleichungen $i<i_0, k<k_0$ erfüllt, also entweder a_i oder b_k Null und daher auch $c_l=0$. Für $l=l_0$ hingegen ist $c_{l_0}=a_{i_0}b_{k_0}\neq 0$; also ist $c\neq 0$ und zwar $f(c)=l_0$ oder $f(ab)=f(a)+f(b)$. Auch hier also, wie bei der Multiplikation gewöhnlicher Zahlen, verschwindet ein Produkt dann und nur dann, wenn ein Faktor verschwindet. Daher zieht, unter Beachtung des distributiven Gesetzes, die Gleichung $ab=ab'$ für $a\neq 0$ die Gleichung $b=b'$ nach sich, und die Gleichung $ab=c$ kann. bei gegebenem $a\neq 0$ und c, höchstens eine Auflösung $b=\frac{c}{a}$ haben.

Die oben erwähnten Einheitskomplexe (S. 203) befolgen die Regel $e_ie_k=e_{i+k}$, als ob sie Potenzen einer festen Basis mit den Indices als Exponenten wären. Der zum Index 0 (der Null des Größensystems J) gehörige Komplex e_0, die „Haupteinheit“, spielt die Rolle der Zahl 1, indem stets $ae_0=e_0a=a$ ist.

Wir haben noch zu zeigen, daß die Gleichung $ab=c$ für $a\neq 0$ stets eine (und nur eine) Auflösung b hat. Ist $f(a)=i$, also $a=a_ie_i+\ldots$ (wobei unter $p+\ldots$ die Summe $p+q$ aus p und einem Komplex von niederer Klasse, also mit $f(q)>f(p)$, verstanden werden soll), so multiplizieren wir die Gleichung $ab=c$ mit $\frac{1}{a_i}e_{-i}$, wodurch sie in eine Gleichung $a'b=c'$ mit $a'=e_0+\ldots$ übergeht. Die zweite Gleichung zieht umgekehrt die erste nach sich. Wir können also gleich von vornherein $f(a)=0$ und $a=e_0+\ldots$ annehmen.

Betrachten wir nun die Menge aller Komplexe y, für die $c-ay$ von Null verschieden ist, und die Menge der zugehörigen Indices $f(c-ay)$ oder Klassen. Wieder ist zu unterscheiden, ob unter diesen Klassen eine niedrigste ist oder nicht, also unter den Indices ein größter oder nicht. Ist eine niedrigste Klasse K_i vorhanden und y ein Komplex, für den

$$f(c-ay)=i,$$
$$c-ay \quad = z_i e_i + \ldots,$$

so folgt aus

$$a \cdot z_i e_i = z_i e_i + \ldots,$$

daß $c-a(y+z_i e_i)$ von niederer Klasse als e_i oder aber Null ist. Da das erste der über y gemachten Annahme widerspricht, so muß das zweite eintreten, also $a(y+z_i e_i)=c$ oder $b=y+z_i e_i$ die Auflösung unserer Gleichung.

Ist keine niedrigste unter den genannten Klassen vorhanden, so ist die Menge K der Indices $k=f(c-ay)$ ohne letztes Element und es sei wieder H die Menge aller Vorgänger von Elementen k (vgl. S. 204). Ist h ein Element von H, so gibt es ein y, wofür $f(c-ay)>h$, und wieder behaupten wir, daß alle diese y an der Stelle h ein festes Element b_h tragen. Denn sind y und y' verschieden und $f(c-ay)$ und $f(c-ay')$ beide $>h$, dann ist nach genau demselben Schluß wie oben für die Differenz beider Komplexe $f\big(a(y'-y)\big)>h$, und da die linke Seite $=f(a)+f(y'-y)=f(y'-y)$, so ist $f(y'-y)>h$, $y_h'=y_h$. Setzen wir diesen gemeinsamen Wert $=b_h$, so erhalten wir durch Sammlung dieser Elemente und Hinzufügung von Nullen für die Indices von $J-H$ wieder einen Komplex b, dessen Kongruenz mit 0 wie oben zu beweisen ist. Wir behaupten, daß dieser Komplex b die Gleichung $c=ab$ erfüllt. Andernfalls wäre nämlich b selbst ein y, und $f(c-ab)=h$ ein Element von H. Es gäbe dann ein y, für das $f(c-ay)>h$, und da $c-ay$ unendlich klein gegen $c-ab$ ist, so würde die Differenz beider von derselben Klasse wie $c-ab$ sein, d. h. $f\big(a(y-b)\big)=h$ und $f(y-b)=h$, also $y_h \neq b_h$, während doch, nach der Bestimmung von b_h, aus $f(c-ay)>h$ $y_h=b_h$ folgen sollte. Man kann das auch so wenden, daß man zeigt: $c-ab$ müßte, wenn $\neq 0$, von kleinerer Klasse als alle $c-ay$ und doch b selbst ein y sein.

Als Beispiel für Größensysteme mit Multiplikation empfehlen wir dem Leser, sich den einfachsten Fall klar zu machen, daß das Argument J die Menge der ganzen Zahlen ..., -2, -1, 0, 1, 2, ... ist; unsere Komplexe $a \neq 0$ haben also die Form

$$a=(\ldots, 0, 0, a_i, a_{i+1}, a_{i+2}, \ldots) \qquad (a_i \neq 0)$$

mit höchstens einer endlichen Zahl von nichtverschwindenden

Elementen mit negativen Indices, und verhalten sich hinsichtlich der vier Spezies wie Potenzreihen

$$a_i x^i + a_{i+1} x^{i+1} + a_{i+2} x^{i+2} + \dots .$$

Wählt man als Argument die Menge der rationalen oder reellen Zahlen, so verhalten sich die Komplexe wie Potenzreihen mit rationalen oder reellen Exponenten, wobei die Glieder mit nichtverschwindenden Koeffizienten aber immer, der Größe der Exponenten nach geordnet, eine wohlgeordnete Menge bilden (die in diesem Fall endlich oder abzählbar ist). Wählt man als Argument eine lexikographisch geordnete Menge von Zahlenpaaren, so verhalten sich die Komplexe wie Potenzreihen zweier Variabler x, y; und so kann man fortfahren. Nichtarchimedische Größensysteme dieser Art sind von G. Veronese, T. Levi-Civitá, D. Hilbert u. a. namentlich zu Untersuchungen über die Grundlagen der Geometrie vielfach aufgestellt worden.

Siebentes Kapitel.

Punktmengen in allgemeinen Räumen.

§ 1. Umgebungen.

In der Anwendung auf die Punktmengen des Raumes, in der Klärung und Verschärfung der geometrischen Grundbegriffe hat die Mengenlehre ihre schönsten Triumphe gefeiert, die selbst von denjenigen zugestanden werden, die sich der abstrakten Mengenlehre gegenüber skeptisch verhalten.

Wir haben uns zunächst über die Stellung der Punktmengentheorie im System der allgemeinen Mengenlehre klar zu werden. Man kann eine Menge rein als System ihrer Elemente behandeln, ohne daß Beziehungen zwischen diesen Elementen in Frage kommen; das ist der Standpunkt, den wir in den ersten drei Kapiteln dieses Buches eingenommen haben. Zweitens aber kann man Relationen zwischen den Elementen in Betracht ziehen, und dafür gibt die Theorie der geordneten Mengen, der die drei folgenden Kapitel gewidmet waren, ein fundamentales Beispiel. Hier handelte es sich, für je zwei Elemente, um eine der drei Beziehungen $a \gtreqless b$, und wir konnten das so auffassen, daß eine Funktion $f(a, b)$ der (geordneten) Paare der Menge gegeben sei, eine Funktion, die allerdings nur drei (bei Beschränkung auf Paare verschiedener Elemente nur zwei)

Werte annehmen konnte. Bei den teilweise geordneten Mengen kam noch eine vierte Relation oder ein vierter möglicher Funktionswert hinzu. Nun steht einer Verallgemeinerung dieser Vorstellung nichts im Wege, und wir können uns denken, daß eine beliebige Funktion der Paare einer Menge definiert, d. h. jedem Paar (a, b) von Elementen einer Menge M ein bestimmtes Element $n = f(a, b)$ einer zweiten Menge N zugeordnet sei. In noch weiterer Verallgemeinerung können wir eine Funktion der Elementtripel, Elementfolgen, Elementkomplexe, Teilmengen u. dgl. von M in Betracht ziehen. Eine ganz allgemein gehaltene Theorie dieser Art würde natürlich erhebliche Komplikationen bedingen und wenig positive Ausbeute liefern. Zu den speziellen Beispielen aber, die ein erhöhtes Interesse beanspruchen, gehört neben der Theorie der geordneten Mengen gerade die Lehre von den Punktmengen im Raume, und zwar ist die grundlegende Beziehung hier wieder eine Funktion der Elementpaare, nämlich die Entfernung zweier Punkte: eine Funktion, die aber jetzt unendlich vieler Werte fähig ist. So angesehen, subsumiert sich die Theorie der Punktmengen mitsamt der Theorie der geordneten Mengen unter das allgemeine Schema einer Lehre von solchen Mengen, in denen binäre Relationen, Relationen zwischen je zwei Elementen der Menge, definiert sind.

Andererseits ist dies nicht die einzig mögliche Auffassung. Auf Grund des Begriffs Entfernung läßt sich z. B. der Begriff einer konvergenten Punktfolge und ihres Limes definieren, und diesen Begriff kann man wieder, mit Ausschaltung des Begriffs Entfernung, zum Fundament der Punktmengentheorie wählen. Es würde sich dann formal um eine Menge M handeln, in der eine Funktion $f(a_1, a_2, \ldots, a_n, \ldots)$ der Elementfolgen definiert, nämlich gewissen Folgen (den konvergenten) ein Element von M selbst (der Limes) zugeordnet ist.

Drittens lassen sich auf Grund der Entfernung jedem Punkt gewisse Teilmengen des Raumes zuordnen, die wir Umgebungen dieses Punktes nennen; und wieder läßt sich dieses System der Umgebungen zur Grundlage der ganzen Theorie machen, mit Elimination des Begriffs Entfernung. Hier wird also eine Menge M unter dem Gesichtspunkt einer Zuordnung zwischen Elementen und Teilmengen betrachtet; wir haben übrigens gezeigt (Kap. IV, § 1), daß man auch die Ordnung einer Menge durch ein passendes System von Teilmengen definieren kann.

Eine Theorie der räumlichen Punktmengen würde nun, vermöge der zahlreichen mitspielenden Eigenschaften des gewöhnlichen Raumes, natürlich einen sehr speziellen Charakter tragen, und wenn man sich von vornherein auf diesen einzigen Fall festlegen wollte,

so würde man für Punktmengen einer Geraden, einer Ebene, einer Kugel usw. jedesmal eine neue Theorie zu entwickeln haben. Die Erfahrung hat gezeigt, daß man diesen Pleonasmus vermeiden und eine allgemeinere Theorie aufstellen kann, die nicht nur die genannten Fälle, sondern auch noch andere Mengen (Riemannsche Flächen, Räume von endlich und unendlich vielen Dimensionen, Kurven- und Funktionenmengen u. a.) umfaßt. Und zwar ist dieser Gewinn an Allgemeinheit nicht etwa mit einer erhöhten Komplikation, sondern gerade umgekehrt mit einer erheblichen Vereinfachung verbunden, indem wir, wenigstens für die Grundzüge der Theorie, nur von ganz wenigen und einfachen Voraussetzungen (Axiomen) Gebrauch zu machen haben. Endlich sichern wir uns auf diesem logisch-deduktiven Wege vor den Irrtümern, zu denen die sogenannte Anschauung uns verleiten möchte; diese angebliche Erkenntnisquelle — deren heuristischen Wert natürlich niemand bestreiten wird — hat sich gerade in den subtileren Fragen der Mengenlehre so oft als unzureichend und unzuverlässig erwiesen, daß man ihren scheinbaren Evidenzen nur nach vorsichtiger Prüfung trauen darf.

Welchen der drei oben genannten Grundbegriffe Entfernung, Limes, Umgebung man zur Basis der Betrachtung wählen will, ist bis zu einem gewissen Grade Geschmacksache. Mit Hilfe von Entfernungen kann man, wie gesagt, Umgebungen und Limites definieren, mit Hilfe von Umgebungen Limites, aber im allgemeinen nicht Entfernungen, mit Hilfe von Limites im allgemeinen weder Umgebungen noch Entfernungen. Danach scheint die Entfernungstheorie die speziellste, die Limestheorie die allgemeinste zu sein; auf der andern Seite bringt der Limesbegriff sofort eine Beziehung zum Abzählbaren (zu Elementfolgen) in die Theorie hinein, worauf die Umgebungstheorie verzichtet. Wir ziehen aus verschiedenen Gründen vor, die grundlegenden Betrachtungen dieses Kapitels auf die Umgebungen zu stützen und die beiden andern Begriffe erst später zur Mitwirkung heranzuziehen; um aber dem Leser sogleich ein konkretes Bild zu erwecken, beginnen wir mit den speziellen Umgebungen, die durch Entfernungen definiert sind.

Unter einem metrischen Raume verstehen wir eine Menge E, in der je zwei Elementen (Punkten) x, y eine reelle nichtnegative Zahl, ihre Entfernung $\overline{xy} \geqq 0$ zugeordnet ist; und zwar verlangen wir überdies die Gültigkeit der folgenden

Entfernungsaxiome:

(α) **(Symmetrieaxiom). Es ist stets** $\overline{yx} = \overline{xy}$.

(β) **(Koinzidenzaxiom). Es ist** $\overline{xy} = 0$ **dann und nur dann, wenn** $x = y$.

(γ) **(Dreiecksaxiom). Es ist stets** $\overline{xy} + \overline{yz} \geqq \overline{xz}$.

14*

Speziell bezeichnen wir als euklidischen n-dimensionalen Zahlenraum E_n die Menge der Komplexe reeller Zahlen

$$x = (x_1, x_2, \ldots, x_n),$$

worin die Entfernung durch

$$\overline{xy} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} \geqq 0$$

definiert ist, und als euklidischen n-dimensionalen Raum einen metrischen Raum, dessen Elemente $\xi, \eta, \ldots$ sich den eben genannten Zahlenkomplexen $x, y, \ldots$ umkehrbar eindeutig und entfernungstreu ($\overline{\xi\eta} = \overline{xy}$) zuordnen lassen. Man nennt dann $x_1, x_2, \ldots, x_n$ die rechtwinkligen Koordinaten des Punktes ξ. Der Leser weiß aus den Elementen der analytischen Geometrie, daß eine solche Abbildung, wenn überhaupt, auf unendlich viele Weisen möglich und wie der Übergang von einer zur andern (Koordinatentransformation) zu bewerkstelligen ist. Er weiß ferner, daß den niedrigsten Fällen ($n = 1, 2, 3$) gewisse Objekte der Anschauung (Gerade, Ebene, Raum) entsprechen, und daß man den hierdurch vermittelten Parallelismus zwischen arithmetischen und geometrischen Tatsachen zu einer auch mehrdimensionale Räume umfassenden geometrischen Ausdrucksweise weitergebildet hat. Wir werden uns übrigens im Fall euklidischer Räume häufig an das anschauliche Vorbild der Ebene halten. Die meisten Tatsachen der Punktmengenlehre in euklidischen Räumen gelten für alle Dimensionenzahlen unterschiedslos; auf gegenteilige Fälle, wo z. B. die eindimensionalen (linearen) Punktmengen eine Sonderstellung einnehmen, ist ausdrücklich hingewiesen.

Denken wir uns für einen euklidischen Raum eine bestimmte Abbildung (Koordinatenwahl) festgehalten, so werden wir von der üblichen Freiheit Gebrauch machen, den Punkt ξ mit seinem Bildkomplex x oder den euklidischen Raum mit dem Zahlenraum einfach zu identifizieren. Gewisse Begriffe, so nach Definition der Begriff Entfernung und der sogleich einzuführende Begriff Umgebung, sind vom Koordinatensystem unabhängig; andere sind es nicht, z. B. die häufig zu benutzende Scheidung in rationale Punkte, deren sämtliche Koordinaten rational sind, und irrationale Punkte mit mindestens einer irrationalen Koordinate. Die Menge der rationalen Punkte hat die Mächtigkeit $\aleph_0^n = \aleph_0$, ist also abzählbar.

In einem metrischen Raume E verstehen wir unter einer Umgebung U_x des Punktes x die Menge der Punkte y, deren Entfernung von x kleiner als eine bestimmte positive Zahl ϱ ist ($\overline{xy} < \varrho$). Eine solche Umgebung hängt vom Mittelpunkt x und vom Radius ϱ ab; ein Punkt x hat, wenn man den Radius variiert,

unendlich viele Umgebungen, die aber als Punktmengen im allgemeinen nicht alle verschieden zu sein brauchen. Auf der geraden Linie E_1 ist U_x eine Strecke (ohne Endpunkte) mit dem Mittelpunkt x und der Länge 2ϱ; in der Ebene E_2 ist U_x das Innere eines Kreises (ohne Peripherie) mit dem Mittelpunkt x und dem Radius ϱ; im Raume E_3 ebenso das Innere einer Kugel, und diesen Namen überträgt man auch auf den Fall eines n-dimensionalen euklidischen und eines metrischen Raumes.

Diese sphärischen Umgebungen, wie wir sie nennen wollen, haben nun eine Reihe von Eigenschaften, von denen wir zunächst nur ganz wenige brauchen. Dabei ändern wir, wie vorhin angekündigt, unseren Standpunkt dahin, daß wir von den Entfernungen, mit deren Hilfe wir Umgebungen definiert haben, absehen und die genannten Eigenschaften demgemäß als Axiome an die Spitze stellen.

Unter einem topologischen[1] Raum verstehen wir eine Menge E, worin den Elementen (Punkten) x gewisse Teilmengen U_x zugeordnet sind, die wir Umgebungen von x nennen, und zwar nach Maßgabe der folgenden

Umgebungsaxiome:

(A) **Jedem Punkt x entspricht mindestens eine Umgebung U_x; jede Umgebung U_x enthält den Punkt x.**

(B) **Sind U_x, V_x zwei Umgebungen desselben Punktes x, so gibt es eine Umgebung W_x, die Teilmenge von beiden ist** $\left(W_x \leqq \mathfrak{D}(U_x, V_x)\right)$.

(C) **Liegt der Punkt y in U_x, so gibt es eine Umgebung U_y, die Teilmenge von U_x ist** $(U_y \leqq U_x)$.

(D) **Für zwei verschiedene Punkte x, y gibt es zwei Umgebungen U_x, U_y ohne gemeinsamen Punkt** $\left(\mathfrak{D}(U_x, U_y) = 0\right)$.

Der Nachweis, daß die sphärischen Umgebungen in einem metrischen Raume diese Axiome erfüllen, ist sehr einfach; wir wollen ihn dennoch ausführen.

(A) Die Menge U_x der Punkte y, für die $\overline{xy} < \varrho$, enthält jedenfalls den Punkt x. Sie braucht übrigens wirklich keinen weiteren Punkt zu enthalten (x kann ein „isolierter Punkt" von E sein, § 3) und zwei Umgebungen mit verschiedenen Radien brauchen nicht verschieden zu sein.

(B) Von zwei Umgebungen mit verschiedenen Radien ist die mit kleinerem Radius Teilmenge der andern. In jedem Falle kann für W_x eine der beiden Mengen U_x, V_x selbst genommen werden.

[1] Der Ausdruck ist in einem verwandten Sinne bereits üblich; wir wollen damit andeuten, daß es sich um Dinge handelt, die ohne Maß und Zahl ausdrückbar sind.

(C) Hat U_x den Radius ϱ und liegt y in U_x ($\overline{xy} < \varrho$), so wähle man die positive Zahl $\sigma \leqq \varrho - \overline{xy}$. Die Umgebung U_y mit dem Radius σ ist dann in U_x enthalten, d. h. aus $\overline{yz} < \sigma$ folgt $\overline{xz} < \varrho$, weil nach dem Dreiecksaxiom

$$\overline{xz} \leqq \overline{xy} + \overline{yz} < \overline{xy} + \sigma \leqq \varrho.$$

Hierbei ist wesentlich, daß U_x durch $\overline{xy} < \varrho$ (nicht $\leqq \varrho$), also z. B. im Fall der euklidischen Ebene als Kreisfläche ohne Peripherie definiert ist.

(D) Ist $x \neq y$, also $\overline{xy} > 0$, so wähle man die positive Zahl $\varrho \leqq \frac{1}{2}\overline{xy}$. Die Umgebungen U_x, U_y mit den Radien ϱ haben dann keinen Punkt gemein, denn für einen solchen Punkt z wäre

$$\overline{xy} \leqq \overline{xz} + \overline{yz} < 2\varrho$$

im Widerspruch zur Wahl von ϱ.

Damit ist die Behauptung bewiesen und ein metrischer Raum als spezieller topologischer Raum erkannt. Wir werden uns später mehrfach überzeugen, daß die Gültigkeit der Umgebungsaxiome keineswegs ein Privileg der sphärischen Umgebungen ist. Hier genüge ein Beispiel, das zugleich zeigt, wie man die geordneten Mengen nach demselben Formalismus wie die Punktmengen behandeln kann: ist E eine geordnete, der Einfachheit wegen offene Menge (d. h. ohne erstes und letztes Element), so verstehe man unter einer Umgebung von x jede das Element x enthaltende Mittelstrecke E_a^b, d. h. für $a < x < b$ die Menge der Elemente y, für die $a < y < b$. Diese Umgebungen erfüllen unsere Axiome, während sich Entfernungen im allgemeinen nicht definieren lassen.

In diesem Kapitel setzen wir also nur die Umgebungsaxiome, d. h. einen topologischen Raum voraus. Die Beispiele allerdings, mit denen wir die allgemeine Theorie illustrieren und die wir, um einer Vermischung vorzubeugen, durch kleineren Druck auszeichnen, sind dem euklidischen Raume resp. der Ebene oder der geraden Linie entnommen.

§ 2. Innere Punkte und Randpunkte.

Ist jetzt A eine Punktmenge, d. h. eine Teilmenge von E, so nennen wir einen Punkt x einen inneren Punkt von A, wenn es eine zu A gehörige Umgebung U_x gibt. Auf Grund des Umgebungsaxioms (A) gehört dann x selbst zu A. Ein Punkt von A, der kein innerer Punkt ist, heißt ein Randpunkt von A. Die Menge A_i der inneren Punkte von A wird auch das Innere von A, die Menge A_r der Randpunkte der Rand von A genannt, und es gilt die Zerlegung in Inneres und Rand

(1) $$A = A_i + A_r.$$

Beispiele: A sei die von einem Kreise in der Ebene E eingeschlossene Fläche, die Peripherie mitgerechnet. Dann sind die Peripheriepunkte Randpunkte, die übrigen innere Punkte.

A sei die Menge der rationalen Punkte der Ebene E (vgl. S. 212). Da jedes U_x auch irrationale Punkte enthält, so hat A gar keine inneren Punkte; es ist $A_i = 0$, $A_r = A$.

A sei die Halbebene über der Abszissenachse ($y > 0$). Jeder Punkt ist innerer Punkt, also $A_i = A$, $A_r = 0$.

Die Beispiele zeigen, daß A_i oder A_r auch Null sein können. Eine Menge ohne Randpunkte, deren sämtliche Punkte also innere Punkte sind, nennen wir ein Gebiet.[1] Ein Gebiet A ist also durch $A_r = 0$ oder $A_i = A$ definiert. Umgekehrt heiße eine Menge ohne innere Punkte eine Randmenge ($A_i = 0$, $A_r = A$). Die Nullmenge rechnen wir sowohl zu den Gebieten wie zu den Randmengen ($A_i = A_r = A = 0$).

Beispiele. Die obere Halbebene ($y > 0$) ist ein Gebiet, desgleichen das Innere eines Kreises, eines Quadrats, einer Ellipse, ohne die Punkte des Umfanges. Die Peripherie eines Kreises, die Menge der rationalen Punkte, die Menge der irrationalen Punkte sind Randmengen.

Der ganze Raum E und jede Umgebung U_x ist ein Gebiet. Das ist unmittelbare Konsequenz der Axiome (A) und (C).

Wenn $A \leqq B$, so ist $A_i \leqq B_i$; denn ein innerer Punkt von A ist auch innerer Punkt von B ($U_x \leqq A \leqq B$).[2] Wir bezeichnen diese einfache Eigenschaft, weil wir uns öfter auf sie berufen müssen, als Kogredienz (von A_i mit A); d. h. wenn den Mengen $A, B, \ldots$ eines Mengensystems andere Mengen $A', B', \ldots$ zugeordnet sind und dabei aus $A \leqq B$ auch $A' \leqq B'$ folgt, so nennen wir A' mit A kogredient.

Jede Teilmenge einer Randmenge ist wieder eine Randmenge (aus $A \leqq B$, $B_i = 0$ folgt $A_i = 0$).

Iteration. Wenn wir die Zerlegung in Inneres und Rand wiederholen und dies durch doppelte Indices andeuten, so folgt

$$(2) \qquad \begin{cases} A_{ii} = A_i, & A_{ir} = 0, \\ A_{ri} = 0, & A_{rr} = A_r \end{cases}$$

oder in Worten: das Innere einer beliebigen Menge ist ein

[1] Dies deckt sich nicht mit dem üblichen durch Weierstrass eingeführten Sprachgebrauch, der von einem Gebiet auch noch Zusammenhang (§ 7) fordert. Ein Weierstrasssches Gebiet ist für uns ein zusammenhängendes Gebiet, und unser Gebiet würde nach der gebräuchlichen Terminologie im allgemeinen eine Gebietsmenge oder Summe von Gebieten sein. Der Begriff einer Menge ohne Randpunkte ist aber so fundamental, daß er entschieden ein eigenes Substantiv verdient.

[2] Natürlich können aber auch Randpunkte von A innere Punkte von B sein.

Gebiet, der Rand einer beliebigen Menge ist eine Randmenge.

Denn ist x innerer Punkt von A, so gibt es eine Umgebung $U_x = U \leqq A$; da U ein Gebiet ist, so ist $U = U_i \leqq A_i$ (wegen der Kogredienz), also x innerer Punkt von A_i. Demnach ist $A_i \leqq A_{ii} \leqq A_i$, womit $A_{ii} = A_i$ folgt. — Ferner folgt aus $A_r \leqq A$ wegen der Kogredienz $A_{ri} \leqq A_i$, andererseits $A_{ri} \leqq A_r$, also $A_{ri} \leqq \mathfrak{D}(A_i, A_r) = 0$.

Verhalten zu Summe und Durchschnitt. Sind $A, B, \ldots$ die Mengen eines beliebigen, endlichen oder unendlichen Mengenkomplexes,

$$S = \mathfrak{S}(A, B, \ldots),\ D = \mathfrak{D}(A, B, \ldots)$$

ihre Summe und ihr Durchschnitt, so folgt aus der Kogredienz unmittelbar

(3) $$S_i \geqq \mathfrak{S}(A_i, B_i, \ldots),\ D_i \leqq \mathfrak{D}(A_i, B_i, \ldots).$$

Darüber hinaus gilt für zwei (oder endlich viele) Mengen und ihren Durchschnitt

$$D = \mathfrak{D}(A, B):$$

(4) $$D_i = \mathfrak{D}(A_i, B_i).$$

In der Tat: ein Punkt x, der zu $\mathfrak{D}(A_i, B_i)$ gehört, hat eine Umgebung $U_x \leqq A$ und eine Umgebung $V_x \leqq B$. Nach dem Umgebungsaxiom (B) gibt es eine Umgebung W_x, die in $\mathfrak{D}(U_x, V_x)$, also in D als Teilmenge enthalten ist, folglich ist x innerer Punkt von D. Demnach ist

$$\mathfrak{D}(A_i, B_i) \leqq D_i$$

und in Verbindung mit der zweiten Formel (3) folgt das Gleichheitszeichen.

I. Die Summe beliebig vieler und der Durchschnitt endlich vieler Gebiete ist wieder ein Gebiet.

Ist nämlich $A_i = A$, $B_i = B$, ..., so folgt aus der ersten Formel (3) $S_i \geqq S$, und da allgemein $S_i \leqq S$, so ist $S_i = S$, also auch S ein Gebiet. Aus (4) folgt ferner $D_i = D$, also ist im Falle endlich vieler Mengen auch D ein Gebiet.

Daß der Durchschnitt beliebig vieler Randmengen wieder eine Randmenge ist, ist trivial, denn jede Teilmenge einer Randmenge ist ja schon eine Randmenge.

Neben (3) erhalten wir auch entsprechende Formeln für die Ränder, nämlich

(5) $$S_r \leqq \mathfrak{S}(A_r, B_r, \ldots,),\ D_r \geqq \mathfrak{D}(A_r, B_r, \ldots).$$

Denn: ein Punkt von S_r gehört zu S, also zu mindestens einem Summanden, etwa zu A; er kann aber nicht zu A_i gehören, da er sonst zu S_i gehören würde, gehört also zu A_r und zu $\mathfrak{S}(A_r, B_r, \ldots)$. Ferner: ein Punkt von $\mathfrak{D}(A_r, B_r, \ldots)$ gehört zu D, aber nicht zu D_i, da er sonst Punkt von A_i und nicht von A_r wäre, also gehört er zu D_r.

Der Durchschnitt aller Umgebungen von x besteht allein aus dem Punkt x,

(6) $$\{x\} = \mathfrak{D}(U_x, V_x, \ldots),$$

wie auf Grund des Axioms (D) leicht nachzuweisen ist. Auch ein geeignetes Teilsystem von solchen Umgebungen (z. B. in einem metrischen Raume die mit den Radien $1, \frac{1}{2}, \frac{1}{3}, \ldots$) kann dieselbe Eigenschaft haben.

Nach I ist die Summe beliebig vieler Umgebungen ein Gebiet. Umgekehrt läßt sich aber auch jedes von Null verschiedene Gebiet als Summe von Umgebungen darstellen. Ist nämlich G ein solches Gebiet und sind $U_x, V_y, \ldots$ alle in G enthaltenen Umgebungen ($U_x \leqq G$, $V_y \leqq G, \ldots$), so ist

(7) $$G = \mathfrak{S}(U_x, V_y, \ldots).$$

In der Tat ist, wenn die Menge rechts mit $\mathfrak{S}$ bezeichnet wird, zunächst $\mathfrak{S} \leqq G$. Andererseits ist aber auch $G \leqq \mathfrak{S}$; denn ist x ein Punkt von G, so gibt es eine Umgebung $U_x \leqq G$, und nach Definition ist dann $U_x \leqq \mathfrak{S}$, x ein Punkt von $\mathfrak{S}$. Hiermit ist die obige Gleichung bewiesen; auch ein geeignetes Teilsystem der in G enthaltenen Umgebungen kann dieselbe Eigenschaft haben.

Die Menge der inneren Punkte beliebig vieler Kreise, gleichviel wie diese Kreise zueinander liegen, ist also ein Gebiet und zwar das allgemeinste Gebiet (der Ebene). Nach I ist, wenn zwei Kreise einander schneiden, auch das Innere des entstehenden Kreisbogenzweiecks ein Gebiet.

Daß eine der Formel (4) entsprechende Formel für $S = \mathfrak{S}(A, B)$ nicht gilt, lehrt ein Beispiel wie dieses: zwei mit einer Seite aneinandergelegte Quadrate A, B (die Umfänge mitgerechnet) bilden ein Rechteck S, dessen Inneres außer den inneren Punkten von A und B auch noch die Punkte der gemeinsamen Seite, bis auf deren Endpunkte, enthält.

Die Formel (6) zeigt, daß der Durchschnitt unendlich vieler, und zwar schon abzählbar vieler Gebiete kein Gebiet zu sein braucht.

Die Summe auch nur zweier Randmengen braucht keine Randmenge zu sein. Die Menge der rationalen und die der irrationalen Punkte geben als Summe das Gebiet E.

Wir machen im Anschluß an den Satz I noch eine einfache Bemerkung, auf die wir im folgenden wiederholt zurückkommen werden. Wenn ein System von Mengen M (wie hier das System der Gebiete) die Eigenschaft hat, daß die Summe beliebig vieler Mengen des Systems wieder dem System angehört, so läßt sich in bezug auf eine beliebige Menge A, die mindestens ein M als Teilmenge enthält, die größte in A enthaltene Menge M definieren. Die Summe aller Mengen $M \leqq A$ ist nämlich selbst ein M, das Teilmenge von A ist, und enthält jedes solche M seinerseits als Teil-

menge. So würde man hier das größte in A enthaltene Gebiet M definieren können, indessen ist dieses nichts neues, sondern einfach das Innere von A. Denn aus $M \leqq A$, $M_i = M$ folgt $M \leqq A_i$; andererseits ist A_i selbst ein Gebiet, also $M \geqq A_i$, und demnach $M = A_i$.

Wenn andererseits ein System von Mengen M (wie das der Randmengen) die Eigenschaft hat, daß der Durchschnitt beliebig vieler Mengen des Systems wieder dem System angehört, so läßt sich für eine Menge A, die in mindestens einem M als Teilmenge enthalten ist, die kleinste, A enthaltende Menge M definieren. Denn der Durchschnitt aller Mengen $M \geqq A$ ist selbst eine solche und in jeder andern als Teilmenge enthalten. So kann man, wenn A in einer Randmenge enthalten, also selbst Randmenge ist, die kleinste Randmenge $\geqq A$ definieren, die aber natürlich A selbst ist; ein weniger triviales Beispiel werden wir in den abgeschlossenen Mengen (§ 3) kennen lernen. Die Gebiete haben die Durchschnittseigenschaft nicht; der Durchschnitt aller Gebiete $\geqq A$ ist (wie aus dem Axiom (D) leicht folgt) A selbst, also im allgemeinen kein Gebiet.

Ist A eine Punktmenge, B ihr Komplement, also

$$(8) \qquad \begin{cases} E = A + B \\ \quad = A_i + A_r + B_i + B_r, \end{cases}$$

so werden die inneren Punkte von B gelegentlich auch als äußere Punkte für A oder von A bezeichnet. Die Randpunkte von A und von B werden Grenzpunkte von A (also auch von B) genannt, und die Menge dieser Grenzpunkte

$$(9) \qquad A_g = A_r + B_r = B_g$$

heißt die Grenze von A oder von B. Die Grenze einer Menge ist also die Summe ihres Randes und des Randes der Komplementärmenge.[1] Ein Gebiet z. B. hat keinen (von 0 verschiedenen) Rand, wohl aber im allgemeinen eine Grenze, nämlich den Rand des Komplements.

Beispiele. Eine Kreisperipherie K zerlegt die Ebene in das Innere J und das Äußere L ($E = J + K + L$). Ist $K = K_1 + K_2$ irgend eine Zerlegung von K in zwei fremde Mengen, so sei $A = J + K_1$, also $B = L + K_2$. Hier ist

$$A_i = J, \; A_r = K_1, \; B_i = L, \; B_r = K_2,$$
$$A_g = B_g = K_1 + K_2 = K,$$

die Kreisperipherie also die Grenze beider Mengen.

[1] Der Sprachgebrauch verschiedener Autoren in bezug auf Randpunkte, Grenzpunkte, Häufungspunkte, Verdichtungspunkte ist leider sehr schwankend.

A sei die Menge der rationalen, B die der irrationalen Punkte. Hier ist

$$A_i = 0,\ A_r = A,\ B_i = 0,\ B_r = B,$$
$$A_g = B_g = A + B = E,$$

die Grenze beider Mengen ist die ganze Ebene.

Die Grenze einer Summe oder eines Durchschnitts von Mengen steht mit den Grenzen der einzelnen Mengen im allgemeinen in keinem einfachen Zusammenhang. Handelt es sich indessen um Gebiete $G_1, G_2, \ldots$, deren Summe und Durchschnitt

$$G = \mathfrak{S}(G_1, G_2, \ldots), \quad G' = \mathfrak{D}(G_1, G_2, \ldots)$$

sei (wobei wir im Fall unendlich vieler Gebiete ausdrücklich voraussetzen, daß auch G' ein Gebiet sei), und werden die Komplemente dieser Gebiete mit dem Buchstaben F, die Grenzen mit H bezeichnet, so ist

$$F = \mathfrak{D}(F_1, F_2, \ldots), \quad F' = \mathfrak{S}(F_1, F_2, \ldots)$$

und nach der Formel (5)

(10) $$H \geqq \mathfrak{D}(H_1, H_2, \ldots), \quad H' \leqq \mathfrak{S}(H_1, H_2, \ldots).$$

Eine etwas inhaltreichere Formel für H erhält man im Fall paarweise fremder Gebiete

$$G = G_1 + G_2 + G_3 + \ldots.$$

Hier ist

$$F_1 = F + G_2 + G_3 + \ldots,$$

und die Randpunkte von F_1 sind, da die Punkte von $G_2, G_3, \ldots$ innere Punkte der rechtsstehenden Menge sind, auch Randpunkte von F, also $H_1 \leqq H$ und demnach

(11) $$H \geqq \mathfrak{S}(H_1, H_2, \ldots).$$

§ 3. Die α-, β-, γ-Punkte.

A sei wieder eine beliebige Punktmenge, und x ein Punkt von E. Wir definieren:

x heißt ein α-Punkt (Berührungspunkt) von A, wenn in jeder Umgebung U_x mindestens ein Punkt von A liegt;

x heißt ein β-Punkt (Häufungspunkt) von A, wenn in jeder Umgebung U_x unendlich viele Punkte von A liegen;

x heißt ein γ-Punkt (Verdichtungspunkt) von A, wenn in jeder Umgebung U_x unabzählbar unendlich viele Punkte von A liegen.[1]

---

[1] Im Rahmen einer allgemein gehaltenen Theorie sind dies nur die ersten drei Glieder einer Skala, die der Skala der Alefs entspricht. Man würde also fortfahren können: x heißt ein β-, γ-, δ-, ... Punkt von A, wenn in jeder

Die Mengen dieser Punkte sollen resp. mit

$$A_\alpha, \ A_\beta, \ A_\gamma$$

bezeichnet werden. G. Cantor nennt A_β die Ableitung von A, und neuerdings hat man A_γ als den Kern von A bezeichnet, welchen Ausdruck wir aber in anderem Sinne brauchen werden. Die Menge A_α ist bisher wenig beachtet worden, zum Schaden der formalen Einfachheit.

Beispiele. Ist A die Menge der rationalen Punkte der Ebene E, so ist

$$A_\alpha = A_\beta = E, \ A_\gamma = 0.$$

Ist A die Menge der irrationalen Punkte, so ist

$$A_\alpha = A_\beta = A_\gamma = E.$$

Haben J, K, L die Bedeutung wie S. 218, und ist $A = J + K_1$, so ist

$$A_\alpha = A_\beta = A_\gamma = J + K.$$

Ist A die Menge der ganzzahligen Punkte (Punkte mit ganzzahligen Koordinaten), so ist

$$A_\alpha = A, \ A_\beta = A_\gamma = 0.$$

Ist A die Menge der Punkte, deren Koordinaten die Reziproken ganzer Zahlen sind (also von der Form $\frac{1}{m}$, wo m eine ganze Zahl $\gtreqless 0$ ist), und bedeutet a den Anfangspunkt des Koordinatensystems, so ist

$$A_\alpha = A + \{a\}, \quad A_\beta = \{a\}, \quad A_\gamma = 0.$$

Jeder Punkt der Menge A selbst ist zugleich α-Punkt, da seine Umgebungen ihn selbst enthalten. Andererseits ist klar, daß ein β-Punkt zugleich α-Punkt, ein γ-Punkt zugleich β-Punkt ist. Demnach haben wir

(1) $$A_\alpha \geqq A, \ A_\alpha \geqq A_\beta \geqq A_\gamma.$$

Genauer haben wir außer diesen Ungleichungen noch die Gleichung

(2) $$A_\alpha = \mathfrak{S}(A, A_\beta).$$

Nach (1) ist nämlich $A_\alpha \geqq \mathfrak{S}(A, A_\beta)$; wir haben also zu zeigen, daß umgekehrt auch $A_\alpha \leqq \mathfrak{S}(A, A_\beta)$, daß also ein Punkt x, der weder zu A noch zu A_β gehört, auch nicht zu A_α gehört. Da der Punkt x nicht zu A_β gehören soll, so hat er eine Umgebung U_0, die nur endlich viele Punkte $a_1, a_2, \ldots, a_m$ von A enthält (sollte U_0 keinen Punkt von A enthalten, so wäre die Behauptung schon bewiesen).

Umgebung U_x mindestens $\aleph_0, \aleph_1, \aleph_2, \ldots$ Punkte von A liegen. Im euklidischen Raume würden wir mit der Wahrscheinlichkeit zu rechnen haben, daß alle Mengen A_δ und die folgenden Null sind, so daß diese allgemeine Theorie kein positives Ergebnis vor der speziellen voraus hätte. Dagegen müßte man z. B. für eine Umgebungstheorie der geordneten Mengen (S. 214) auch die Verdichtungspunkte höherer Ordnung in Betracht ziehen.

Da x auch kein Punkt von A sein soll, so ist $a_1 \neq x$, und nach dem Axiom (D) gibt es eine Umgebung U_1 von x, die den Punkt a_1 nicht enthält; analog definieren wir $U_2, \ldots, U_m$. Nach dem Axiom (B) gibt es eine Umgebung

$$U \leqq \mathfrak{D}(U_0, U_1, U_2, \ldots, U_m);$$

dieser Durchschnitt, also auch U, enthält keinen Punkt von A, folglich ist x kein α-Punkt von A, und die Gleichung (2) ist bewiesen.

Setzen wir noch

$$A_h = \mathfrak{D}(A, A_\beta); \tag{3}$$

dies ist also die Menge derjenigen Punkte von A selbst, die zugleich Häufungspunkte sind. Die Vergleichung von (2) und (3) lehrt:

$$A_\alpha - A = A_\beta - A_h$$

ist die Menge derjenigen Häufungspunkte von A, die nicht zu A selbst gehören, und

$$A_\alpha - A_\beta = A - A_h$$

ist die Menge derjenigen Punkte von A, die nicht Häufungspunkte sind. Diese Punkte von A nennt man isolierte Punkte; nach dem Beweis der Formel (2) gibt es zu einem isolierten Punkt a eine Umgebung U_a, die keinen Punkt von A außer a selbst enthält, womit sich der Name erklärt. Bezeichnen wir noch die Menge der isolierten Punkte[1] mit A_j, so ist also

$$A = A_h + A_j, \quad A_\alpha = A_\beta + A_j. \tag{4}$$

Wir gehen nun dazu über, spezielle Kategorien von Mengen nach dem Verhalten von A zu A_β zu definieren (zwischen A und A_α besteht ja immer die Relation $A \leqq A_\alpha$). Im allgemeinen ist keine dieser Mengen Teilmenge der andern, d. h. es gibt Häufungspunkte, die nicht zu A gehören, und Punkte von A, die keine Häufungspunkte (sondern isolierte Punkte) sind. Die Mengen, wo eine der drei Relationen $A \gtreqless A_\beta$ besteht, werden folgendermaßen benannt: die Menge A heißt

abgeschlossen, wenn $A \geqq A_\beta$,
insichdicht[2], wenn $A \leqq A_\beta$,
perfekt, wenn $A = A_\beta$.

Also: die Menge A ist abgeschlossen, wenn sie ihre Häufungspunkte

---

[1] Der Buchstabe i ist schon für innere Punkte vergriffen.

[2] Wir schreiben dies in einem Wort, um einer Verwechselung vorzubeugen: später werden wir eine Menge in einer anderen Menge dicht nennen unter Bedingungen, nach denen jede Menge in sich selber dicht ist. Wenn wir nicht die Abweichungen vom Cantorschen Sprachgebrauch auf das Mindestmaß einschränken wollten, würden wir eine insichdichte Menge lieber eine gehäufte Menge nennen.

enthält; insichdicht, wenn jeder ihrer Punkte Häufungspunkt ist; perfekt, wenn beides der Fall ist. Diese Ungleichungen lassen sich sämtlich durch Gleichungen ersetzen.

Ist die Menge abgeschlossen, $A \geqq A_\beta$, so ist nach (2) $A_\alpha = A$. Ist umgekehrt $A = A_\alpha$, so ist nach (1) $A \geqq A_\beta$.

Ist die Menge insichdicht, $A \leqq A_\beta$, so ist nach (2) $A_\alpha = A_\beta$. Ist umgekehrt $A_\alpha = A_\beta$, so ist nach (1) $A \leqq A_\beta$.

Also können diese Mengen auch durch die Gleichungen definiert werden: A heißt

abgeschlossen, wenn $A = A_\alpha$,
insichdicht, wenn $A_\alpha = A_\beta$,
perfekt, wenn $A = A_\beta (= A_\alpha)$.

Auch mit Hilfe der in (3) definierten Menge A_h können diese Eigenschaften erklärt werden. Für eine abgeschlossene Menge ist $A_h = A_\beta$; für eine insichdichte Menge $A_h = A$, also $A_j = 0$, eine insichdichte Menge hat keine isolierten Punkte. Dies gibt den Anlaß, noch eine vierte spezielle Eigenschaft als Gegenstück zur Insichdichtheit zu definieren: die Menge A heißt isoliert, wenn sie nur aus isolierten Punkten besteht, also mit der Menge ihrer Häufungspunkte keinen Punkt gemein hat $(\mathfrak{D}(A, A_\beta) = 0)$. Um also dieses Begriffspaar noch einmal gegenüber zu stellen, so heißt die Menge A

insichdicht, wenn $A_j = 0$, $A = A_h$,
isoliert, wenn $A_h = 0$, $A = A_j$.

Wir haben wieder ausdrücklich hervorzuheben, daß wir das Verschwinden einer der definierten Mengen nicht ausschließen. So nennen wir eine Menge, zu der überhaupt keine Häufungspunkte existieren ($A_\beta = 0$), insbesondere jede endliche Menge, abgeschlossen und isoliert. Die Nullmenge wäre danach gleichzeitig als abgeschlossen, insichdicht, perfekt und isoliert zu bezeichnen; in § 2 erschien sie auch als Gebiet und als Randmenge.[1]

Beispiele (vgl. S. 220). Die Menge der rationalen wie der irrationalen Punkte ist insichdicht, aber nicht abgeschlossen. Die Menge der ganzzahligen Punkte ist abgeschlossen und isoliert. Die Menge $J + K_1$ (Kreisinneres plus einem Teil der Peripherie K) ist insichdicht, aber für $K_1 \subset K$ nicht abgeschlossen; die Menge $J + K$ ist auch abgeschlossen, also perfekt. Ist a ein Punkt außerhalb des Kreises, so ist die Menge

[1] Die Behandlung der Nullmenge ist natürlich Geschmacksache; wie man sich auch dazu stelle, wird man gelegentliche Schwerfälligkeiten nicht vermeiden können. Dagegen halten wir es unter allen Umständen für unzweckmäßig, etwa die endlichen Mengen nicht mit zu den abgeschlossenen Mengen zu rechnen, wie dies mehrfach üblich ist.

$J + K + \{a\}$ abgeschlossen, aber nicht insichdicht; die Menge ihrer isolierten Punkte ist $\{a\}$.

Kogredienz. Wenn $A \leqq B$, so ist auch

$$A_\alpha \leqq B_\alpha, \quad A_\beta \leqq B_\beta, \quad A_\gamma \leqq B_\gamma,$$

wie unmittelbar aus der Definition folgt, und aus der mittleren Formel und $A \leqq B$ selbst ergibt sich noch vermöge Durchschnittsbildung

$$A_h \leqq B_h.$$

Die Mengen A_α, A_β, A_γ, A_h (nicht A_j) sind also mit A kogredient. Jede Teilmenge einer isolierten Menge ist wieder isoliert.

Iteration. Wendet man auf die erhaltenen Mengen dieselben Prozesse noch einmal an und bezeichnet dies durch Doppelindices (z. B. $A_{\alpha\beta}$ ist die Menge der β-Punkte von A_α), so gilt[1]

(5) $$A_{\alpha\alpha} = A_\alpha, \quad A_{\beta\alpha} = A_\beta, \quad A_{\gamma\alpha} = A_\gamma,$$

(5′) $$A_{\alpha\beta} = A_\beta.$$

Um die Formeln (5) zu beweisen, nehmen wir an, x sei ein α-Punkt von A_α, resp. A_β, resp. A_γ, und zeigen, daß x dann auch Punkt dieser Mengen ist. Jede Umgebung U_x enthält mindestens einen Punkt y von A_α, resp. A_β, resp. A_γ, und nach dem Axiom (C) gibt es eine Umgebung $U_y \leqq U_x$. U_y enthält dann mindestens einen, resp. unendlich viele, resp. unabzählbar viele Punkte von A, und das Gleiche gilt also von U_x, womit x als Punkt von A_α, resp. A_β, resp. A_γ erwiesen ist. Daraus folgt

$$A_{\alpha\alpha} \leqq A_\alpha, \quad A_{\beta\alpha} \leqq A_\beta, \quad A_{\gamma\alpha} \leqq A_\gamma$$

und in Verbindung mit der allgemeinen Relation $M \leqq M_\alpha$ das Gleichheitszeichen. Die Formeln (5) besagen:

Die Mengen A_α, A_β, A_γ sind stets abgeschlossen.

Um (5′) zu beweisen, nehmen wir an, x sei ein β-Punkt von A_α und U_x eine beliebige Umgebung. U_x enthält unendlich viele Punkte von A_α. Sind diese allesamt Punkte von A, so enthält U_x unendlich viele Punkte von A. Gehört einer von ihnen, y, nicht zu A, so gehört er nach (2) zu A_β; es gibt dann eine Umgebung $U_y \leqq U_x$, die unendlich viele Punkte von A enthält. Auch in diesem, also in jedem Fall enthält U_x unendlich viele Punkte von A, x ist ein β-Punkt von A. Danach ist $A_{\alpha\beta} \leqq A_\beta$, andererseits wegen der Kogredienz (weil $A_\alpha \geqq A$) auch $A_{\alpha\beta} \geqq A_\beta$, womit (5′) bewiesen ist.

Die Menge A_β ist also die Menge der Häufungspunkte nicht nur von A, sondern auch von A_α. Ist A insichdicht, so ist A_α perfekt, und umgekehrt.

---

[1] Weitere Formeln dieser Art können wir erst im folgenden Kapitel angeben (S. 270).

Auch für die Spaltung in Häufungspunkte und isolierte Punkte gelten gewisse Iterationsformeln, nämlich

(6) $$A_{jh} = 0, \quad A_{jj} = A_j.$$

Denn wegen $A_j \leqq A$ ist $A_{j\beta} \leqq A_\beta$,

$$A_{jh} = \mathfrak{D}(A_j, A_{j\beta}) \leqq \mathfrak{D}(A_j, A_\beta) = 0.$$

Die Menge A_j der isolierten Punkte einer beliebigen Menge ist also selbst eine isolierte Menge. Dagegen ist nicht etwa $A_{hj} = 0$ und A_h im allgemeinen keine insichdichte Menge; denn die Punkte von A_h sind zwar Häufungspunkte von A, brauchen aber keine Häufungspunkte von A_h (Häufungspunkte von Häufungspunkten) zu sein.

Verhalten zu Summe und Durchschnitt. Sind wieder $A, B, \ldots$ Mengen eines endlichen oder unendlichen Komplexes und

$$S = \mathfrak{S}(A, B, \ldots), \quad D = \mathfrak{D}(A, B, \ldots),$$

so folgt aus der Kogredienz

(7) $$\begin{cases} S_\alpha \geqq \mathfrak{S}(A_\alpha, B_\alpha, \ldots), & D_\alpha \leqq \mathfrak{D}(A_\alpha, B_\alpha, \ldots) \\ S_\beta \geqq \mathfrak{S}(A_\beta, B_\beta, \ldots), & D_\beta \leqq \mathfrak{D}(A_\beta, B_\beta, \ldots) \\ S_\gamma \geqq \mathfrak{S}(A_\gamma, B_\gamma, \ldots), & D_\gamma \leqq \mathfrak{D}(A_\gamma, B_\gamma, \ldots) \end{cases}$$

Für zwei (oder endlich viele) Mengen und ihre Summe $S = \mathfrak{S}(A, B)$ gilt aber das Gleichheitszeichen:

(8) $$S_\alpha = \mathfrak{S}(A_\alpha, B_\alpha), \quad S_\beta = \mathfrak{S}(A_\beta, B_\beta), \quad S_\gamma = \mathfrak{S}(A_\gamma, B_\gamma).$$

Um die erste dieser Gleichungen zu beweisen, haben wir nach (7) zu zeigen, daß $S_\alpha \leqq \mathfrak{S}(A_\alpha, B_\alpha)$ ist, oder daß ein Punkt, der weder zu A_α noch zu B_α gehört, auch nicht zu S_α gehört; und entsprechend für die beiden andern Formeln.

Gehört x weder zu A_α noch zu B_α, so gibt es eine Umgebung U_x, die keinen Punkt von A enthält, und eine Umgebung V_x, die keinen Punkt von B enthält. Der Durchschnitt $\mathfrak{D}(U_x, V_x)$ enthält dann nach dem Axiom (B) wieder eine Umgebung von x als Teilmenge und diese enthält keinen Punkt von S, also gehört x auch nicht zu S_α.

Ersetzt man hierin den Index α durch β, resp. γ, und die gesperrt gedruckten Worte keinen Punkt durch die Worte höchstens endlich viele Punkte resp. höchstens abzählbar viele Punkte, so hat man den Beweis der beiden andern Formeln.

Aus der Kogredienz von A_h mit A folgt noch weiter

(9) $$S_h \geqq \mathfrak{S}(A_h, B_h, \ldots), \quad D_h \leqq \mathfrak{D}(A_h, B_h, \ldots)$$

und daraus (analog wie in § 2 die Formel (5) aus (3) folgt)

(10) $$S_j \leqq \mathfrak{S}(A_j, B_j, \ldots), \quad D_j \geqq \mathfrak{D}(A_j, B_j, \ldots).$$

Aus diesen Summen- und Durchschnittsformeln ergeben sich wieder verschiedene Sätze:

I. Der Durchschnitt beliebig vieler und die Summe endlich vieler abgeschlossener Mengen ist wieder abgeschlossen.

II. Die Summe beliebig vieler insichdichter Mengen ist wieder insichdicht.

Beweis von I. Ist $A_\beta \leqq A$, $B_\beta \leqq B, \ldots$, so folgt aus (7)

$$D_\beta \leqq \mathfrak{D}(A_\beta, B_\beta, \ldots) \leqq \mathfrak{D}(A, B, \ldots) = D,$$

also ist auch D abgeschlossen. Für zwei Mengen (oder endlich viele) folgt aus (8)

$$S_\beta = \mathfrak{S}(A_\beta, B_\beta) \leqq \mathfrak{S}(A, B) = S,$$

also ist auch S abgeschlossen.

Oder: ist $A_\alpha = A$, $B_\alpha = B, \ldots$, so ist nach (7) $D_\alpha \leqq D$, andererseits immer $D_\alpha \geqq D$, also $D_\alpha = D$, und nach (8) $S_\alpha = S$.

Beweis von II. Ist $A_\beta \geqq A$, $B_\beta \geqq B, \ldots$, so ist nach (7)

$$S_\beta \geqq \mathfrak{S}(A_\beta, B_\beta, \ldots) \geqq \mathfrak{S}(A, B, \ldots) = S,$$

also auch S insichdicht.

Oder: ist $A_j = B_j = \ldots = 0$, so ist nach (10) auch $S_j = 0$.

Der Durchschnitt beliebig vieler isolierter Mengen ist wieder isoliert, was aber trivial ist, denn schon jede Teilmenge einer isolierten Menge ist isoliert.

Die Summe unendlich vieler abgeschlossener Mengen braucht nicht abgeschlossen zu sein, nicht einmal die Summe abzählbar vieler. Auf der geraden Linie ist die Menge der Punkte mit den Abszissen $1, \frac{1}{2}, \frac{1}{3}, \ldots$ nicht abgeschlossen; aber sie ist die Summe abzählbar vieler abgeschlossener Mengen, nämlich solcher, die aus je einem Punkt bestehen.

Der Durchschnitt auch nur zweier insichdichter Mengen braucht nicht insichdicht zu sein. Z. B. sind die geraden Linien der Ebene insichdichte Punktmengen; der Durchschnitt zweier sich schneidender Geraden besteht nur aus dem Schnittpunkt, ist also keine insichdichte Menge.

Die Summe auch nur zweier isolierter Mengen braucht nicht isoliert zu sein. Die lineare Punktmenge, bestehend aus den Punkten $1, \frac{1}{2}, \frac{1}{3}, \ldots$, ist isoliert; die Menge, die aus dem einen Punkt 0 besteht, ist isoliert; die Summe beider ist nicht isoliert, da sie den zur Menge gehörigen Häufungspunkt 0 hat.

Nach den Sätzen I, II und mit Rücksicht auf die in § 2 gemachten Bemerkungen kann man folgende Mengen definieren:

Die kleinste abgeschlossene Menge $\geqq A$. Da A stets Teilmenge einer abgeschlossenen Menge (z. B. E) ist, so existiert

diese Menge jedenfalls; sie ist aber nichts Neues, sondern mit A_α identisch. Denn nennt man sie M, so folgt aus $M \geqq A$: $M = M_\alpha \geqq A_\alpha$; andererseits, weil A_α selbst abgeschlossen ist, $M \leqq A_\alpha$, demnach $M = A_\alpha$.

Die größte insichdichte Teilmenge von A. Auch diese existiert stets, da A wenigstens die insichdichte Teilmenge 0 hat. Sie ist von großer Wichtigkeit in der Theorie der Punktmengen; wir nennen sie den Kern von A und bezeichnen sie mit A_k. Ihre Punkte heißen Kernpunkte, die übrigen Punkte von A separierte Punkte, deren Menge A_s der separierte Bestandteil von A. Wir erhalten damit noch eine Zerlegung außer den bisherigen:

(11) $$A = A_k + A_s.$$

Dabei ist der Kern A_k die Summe aller insichdichten Teilmengen von A.

Ist A selbst insichdicht, so ist $A_k = A$, $A_s = 0$.

Enthält A keine insichdichte Teilmenge (außer 0), so ist $A_k = 0$, $A_s = A$. Eine solche Menge heißt separiert.

Offenbar gelten die Iterationsformeln

(12) $$\begin{cases} A_{kk} = A_k, & A_{ks} = 0, \\ A_{sk} = 0, & A_{ss} = A_s. \end{cases}$$

Denn A_k ist insichdicht, also mit seinem Kern identisch. Andererseits ist A_{sk} insichdicht, also $A_k + A_{sk}$ nach II eine insichdichte Teilmenge von A; da A_k aber die größte solche Teilmenge ist, so muß $A_{sk} = 0$ sein. Es folgt daraus, daß der separierte Bestandteil einer beliebigen Menge stets eine separierte Menge ist.

Zu der Spaltung in Häufungspunkte und isolierte Punkte verhält sich die jetzt betrachtete folgendermaßen. Jede insichdichte Teilmenge von A ist auch Teilmenge von A_h, denn aus

$$B \leqq A, \quad B \leqq B_\beta \leqq A_\beta$$

folgt

$$B \leqq \mathfrak{D}(A, A_\beta) = A_h.$$

Der Kern von A ist also auch Kern von A_h, folglich auch von A_{hh}, von A_{hhh} usw.; umgekehrt gehören zu den separierten Punkten von A nicht nur die isolierten (Menge A_j), sondern auch die isolierten Punkte von A_h (A_{hj}), von A_{hh} (A_{hhj}) usw. Wir haben also, wenn wir endliche Iterationen dieser Spaltung in Betracht ziehen:

$$A_k \leqq \mathfrak{D}(A, A_h, A_{hh}, A_{hhh}, \ldots)$$

$$A_s \geqq A_j + A_{hj} + A_{hhj} + A_{hhhj} + \cdots.$$

In Kap. VIII, § 4 werden wir sehen, wie die unendlich fort-

gesetzte Abspaltung der isolierten Punkte schließlich den Kern übrig läßt.[1]

Eine isolierte Menge ist auch separiert, aber nicht umgekehrt.

Beispiel. Auf der Geraden sei A die Menge der Punkte, deren Abszissen von der Form $\frac{1}{p} + \frac{1}{q}$ sind, p und q natürliche Zahlen, von denen wir $p \leqq q$ voraussetzen können. Bezeichnen wir bei festem p die Menge der Zahlen $\frac{1}{p} + \frac{1}{q}$ mit A_p, also

$$A_p = \left\{\frac{1}{p} + \frac{1}{p},\ \frac{1}{p} + \frac{1}{p+1},\ \frac{1}{p} + \frac{1}{p+2},\ \ldots\right\},$$

so ist

$$A = \mathfrak{S}(A_1, A_2, A_3, \ldots).$$

Die Menge A_p hat den Punkt $\frac{1}{p}$ als einzigen Häufungspunkt; diese Punkte $\frac{1}{p}$ und der Punkt 0, der ja Häufungspunkt der Punkte $\frac{1}{p} + \frac{1}{p} = \frac{2}{p}$ ist, sind also Häufungspunkte von A. Weitere Häufungspunkte hat A nicht. Soll nämlich x Häufungspunkt sein, so darf offenbar x nicht negativ sein; $x = 0$ ist Häufungspunkt; für positives x können wir eine natürliche Zahl $p > \frac{2}{x}$ wählen, so daß $\frac{2}{p} < x$. Wir setzen dann

$$B = \mathfrak{S}(A_1, A_2, \ldots, A_{p-1}), \quad C = \mathfrak{S}(A_p, A_{p+1}, \ldots),$$
$$A = \mathfrak{S}(B, C), \quad A_\beta = \mathfrak{S}(B_\beta, C_\beta).$$

Da nun alle Zahlen der Menge A_p höchstens gleich $\frac{2}{p}$ sind und dasselbe von den Zahlen der Menge C gilt, so ist x kein Häufungspunkt von C und muß also Punkt der Menge

$$B_\beta = \left\{1, \frac{1}{2}, \ldots, \frac{1}{p-1}\right\}$$

sein, die aus den (einzigen) Häufungspunkten von $A_1, A_2, \ldots, A_{p-1}$ besteht.

Hiermit ist bewiesen, daß A außer den genannten keine Häufungspunkte hat, also

$$A_\beta = \left\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}.$$

Da die Punkte $\frac{1}{p} = \frac{1}{2p} + \frac{1}{2p}$ zu A gehören, so ist

$$A_h = \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}$$

und diese Menge ist isoliert, da ihr einziger Häufungspunkt 0 nicht zu

[1] G. Cantor nennt A_h die Kohärenz, A_j die Adhärenz von A; A_{hh} die zweite Kohärenz usw., sodaß A_k dann die „letzte Kohärenz" wird.

ihr gehört. Also ist $A_{hh} = 0$ und $A_k = 0$, die Menge A separiert. Man sieht hier, wie die Punkte von A_h zwar Häufungspunkte von A sind, aber in A_h zu isolierten Punkten werden.

Der Leser diskutiere die Menge A der Zahlen $\frac{1}{p} + \frac{1}{q} + \frac{1}{r}$ (p, q, r natürliche Zahlen) und zeige, daß hier erst $A_{hhh} = 0$ wird; auch diese Menge ist also separiert.

Wir haben jetzt noch die Verbindung dieses Paragraphen mit dem vorigen herzustellen. Ist A eine beliebige Menge und B ihr Komplement, also

$$E = A + B,$$

so bestehen für irgend einen Punkt x von E nur zwei Möglichkeiten: entweder gibt es eine Umgebung U_x, die keinen Punkt von A enthält, oder es gibt keine solche. Im ersten Fall gehört diese Umgebung U_x zu B, also ist x innerer Punkt von B; im zweiten enthält jedes U_x mindestens einen Punkt von A, also ist x ein α-Punkt von A. Die Mengen A_α und B_i sind also Komplemente:

(13) $$E = A_\alpha + B_i = A_i + B_\alpha,$$

das Innere einer Menge und die Menge der Berührungspunkte des Komplements sind Komplemente.

Ist insbesondere $A_\alpha = A$, so ist $B_i = B$, d. h. ist A abgeschlossen, so ist B ein Gebiet, und vice versa.

Das Komplement einer abgeschlossenen Menge ist ein Gebiet, und das Komplement eines Gebietes ist eine abgeschlossene Menge.

Von den Sätzen § 2, I und § 3, I kann hiernach der eine aus dem andern vermöge des Prinzips der Dualität (Kap. I, § 4) gefolgert werden. Abgeschlossene Mengen und Gebiete sind die wichtigsten und einfachsten Objekte der Theorie der Punktmengen.

Die Menge E und die Nullmenge sind gleichzeitig abgeschlossen und Gebiete (im euklidischen Raume wird sich zeigen, daß dies die einzigen abgeschlossenen Gebiete sind).

Vergleichen wir (13) noch mit der Formel

$$E = A_i + A_r + B_i + B_r = A + B_i + B_r = A_i + A_r + B,$$

so folgt

$$A_\alpha = A + B_r, \quad B_\alpha = B + A_r;$$

die Menge $A_\alpha - A$ der nicht zu A gehörigen Häufungspunkte von A ist also mit dem Rand B_r des Komplements identisch.

Die Grenze $A_g = A_r + B_r$ einer beliebigen Menge ist stets abgeschlossen, da ihr Komplement $A_i + B_i$ als Summe zweier Gebiete ein Gebiet ist, oder auch, weil $A_g = \mathfrak{D}(A_\alpha, B_\alpha)$ der Durchschnitt abgeschlossener Mengen ist.

Wenn wir die in den beiden vorangehenden Paragraphen aus einer Menge A abgeleiteten Mengen A_i, A_r, A_g, A_α, A_β, A_γ, A_h, A_j, A_k, A_s noch einmal überblicken, so sehen wir, daß alle diese Mengen sich mit Hilfe von Summen-, Durchschnitts- und Differenzenbildung (aus endlich vielen Mengen) auf die drei A_β, A_γ, A_k (Menge der Häufungs-, Verdichtungs-, Kernpunkte) zurückführen lassen. Denn A_α ist die Summe von A und A_β, A_h der Durchschnitt beider, A_j das Komplement $A - A_h$ und A_s das Komplement $A - A_k$; nach den letzten Betrachtungen endlich lassen sich Inneres, Rand und Grenze durch α-Mengen ausdrücken.

Noch weitergehend können wir zeigen, daß sich bei Zulassung von Summen- und Durchschnittsbildung aus unendlich vielen Mengen alle unsere Mengen durch α-Mengen ausdrücken lassen. Es ist nämlich

x ein β-Punkt von A dann und nur dann, wenn er α-Punkt jeder Teilmenge A' ist, die alle Punkte von A bis auf endlich viele enthält ($A - A'$ endlich);

x ein γ-Punkt von A dann und nur dann, wenn er α-Punkt jeder Teilmenge A'' ist, die alle Punkte von A bis auf höchstens abzählbar viele enthält ($A - A''$ höchstens abzählbar).

In der Tat ist wegen

$$(A - A')_\beta = 0,\ (A - A'')_\gamma = 0:$$

$$A_\beta = \mathfrak{S}\left(A'_\beta, (A - A')_\beta\right) = A'_\beta \leqq A'_\alpha,$$

$$A_\gamma = \mathfrak{S}\left(A''_\gamma, (A - A'')_\gamma\right) = A''_\gamma \leqq A''_\beta \leqq A''_\alpha.$$

Ein Punkt x von A_β gehört also allen A'_α an; umgekehrt, ist y kein Häufungspunkt von A, so hat er eine Umgebung U_y, für die $\mathfrak{D}(A, U_y)$ endlich ist, und ist kein α-Punkt von $A' = A - \mathfrak{D}(A, U_y)$, gehört also nicht allen A'_α an. Demnach ist $A_\beta = \mathfrak{D} A'_\alpha$. — Ebenso gehört ein Punkt x von A_γ allen A''_α an; ist umgekehrt y kein γ-Punkt von A, so gibt es ein U_y mit höchstens abzählbarem $\mathfrak{D}(A, U_y)$, und y ist kein α-Punkt von $A'' = A - \mathfrak{D}(A, U_y)$, gehört also nicht allen A''_α an. Folglich ist $A_\gamma = \mathfrak{D} A''_\alpha$; übrigens auch $A_\gamma = \mathfrak{D} A''_\beta$.

Letzten Endes lassen sich also alle unsere Mengen auf Mengen von α-Punkten (oder, wenn man will, auf Mengen von inneren Punkten) zurückführen.

§ 4. Divergente, kompakte, konvergente Mengen.

Nach der Definition der β- und γ-Punkte ist begreiflich, daß unendliche Mengen ohne Häufungspunkt ($A_\beta = 0$) und unabzählbare Mengen ohne Verdichtungspunkt ($A_\gamma = 0$) eine besondere Rolle spielen werden. Von den letzteren wollen wir indessen absehen, da sich

in den meisten von uns betrachteten Räumen ihre Nichtexistenz herausstellen wird (Kap. VIII, § 3, I). Eine unendliche Menge ohne Häufungspunkt nennen wir divergent;[1] das ist ein Spezialfall einer unendlichen isolierten Menge $(\mathfrak{D}(A, A_\beta) = 0)$. Im euklidischen Raume ist z. B. die Menge der ganzzahligen Punkte divergent; die Menge der Punkte, deren Koordinaten die Reziproken ganzer Zahlen sind, ist unendlich und isoliert.

Eine Menge ohne divergente Teilmenge nennen wir (nach M. Fréchet) kompakt; dazu rechnen wir natürlich auch die endlichen Mengen inkl. der Nullmenge. Jede unendliche Teilmenge einer kompakten Menge A hat also mindestens einen (nicht notwendig zu A gehörigen) Häufungspunkt. Jede Teilmenge einer kompakten Menge ist kompakt; die Summe endlich vieler kompakter Mengen ist wieder kompakt.[2]

Wir schließen hieran zwei sehr wichtige Sätze, die sich auf kompakte abgeschlossene Mengen beziehen.[3]

I. (Durchschnittssatz von Cantor). Eine absteigende Folge $A_1 \geqq A_2 \geqq \ldots$ kompakter, abgeschlossener, von Null verschiedener Mengen hat einen von Null verschiedenen Durchschnitt.

Wir wählen aus jeder Menge A_n einen Punkt a_n; die Menge der verschiedenen unter den Punkten $a_n, a_{n+1}, a_{n+2}, \ldots$ (die ja alle in A_n liegen) sei $B_n \leqq A_n$. Ist B_1 endlich, so ist für unendlich viele Indices $a_n = a$; der Punkt a gehört dann zu unendlich vielen, folglich zu allen Mengen A_n, also auch zu ihrem Durchschnitt $D = \mathfrak{D}(A_1, A_2, \ldots)$. Ist aber B_1 unendlich, so hat es als Teilmenge der kompakten Menge A_1 mindestens einen Häufungspunkt a; für jedes n ist dann a auch Häufungspunkt von B_n, also auch von A_n. Wegen der Abgeschlossenheit von A_n gehört dann a zu A_n selbst, also wieder zu D.

Die Voraussetzung der Kompaktheit ist, wie übrigens die andern auch, wesentlich. Auf der geraden Linie ist die Halbgerade A_n, bestehend aus den Punkten, deren Abszisse $x \geqq n$, abgeschlossen, aber nicht kompakt (beschränkt). Alle andern Voraussetzungen sind erfüllt, trotzdem ist $\mathfrak{D}(A_1, A_2, \ldots) = 0$.

---

[1] Nach diesem Sprachgebrauch ist Divergenz nicht das kontradiktorische Gegenteil von Konvergenz (s. u.), sondern nur ein spezieller Fall von Nichtkonvergenz, den viele Autoren als eigentliche Divergenz bezeichnen. Die Menge der rationalen Punkte z. B. ist weder konvergent noch divergent.

[2] Im euklidischen Raume, der selbst nicht kompakt ist, wird sich zeigen, daß kompakte und beschränkte Mengen identisch sind (Kap. VIII, § 10).

[3] Diese Mengen nennt M. Fréchet extremal, weil jede in ihnen definierte stetige reelle Funktion beschränkt ist und Extrema (Maximum und Minimum) hat (Kap. IX, § 1).

II. (Satz von Borel). **Ist eine kompakte abgeschlossene Menge in der Summe einer Folge von Gebieten enthalten, so ist sie bereits in einer Summe von endlich vielen Gebieten dieser Folge enthalten.**

A sei kompakt, abgeschlossen und

$$A \leqq S = \mathfrak{S}(G_1, G_2, \ldots).$$

Die Behauptung lautet, daß bereits für ein gewisses n (und natürlich alle folgenden)

$$A \leqq S_n = \mathfrak{S}(G_1, G_2, \ldots, G_n).$$

Setzen wir

$$B_n = \mathfrak{D}(A, S_n), \quad A_n = A - B_n,$$

so ist

$$S_1 \leqq S_2 \leqq \ldots, \; B_1 \leqq B_2 \leqq \ldots, \; A_1 \geqq A_2 \geqq \ldots,$$

ferner

$$\mathfrak{S}(S_1, S_2, \ldots) = \mathfrak{S}(G_1, G_2, \ldots) = S.$$

daraus durch Schneiden mit A

$$\mathfrak{S}(B_1, B_2, \ldots) = \mathfrak{D}(A, S) = A,$$

also

$$\mathfrak{D}(A_1, A_2, \ldots) = A - \mathfrak{S}(B_1, B_2, \ldots) = 0.$$

Nun ist jedes G_n und S_n ein Gebiet, $E - S_n$ abgeschlossen, $A_n = \mathfrak{D}(A, E - S_n)$ abgeschlossen und kompakt. Wären alle $A_n > 0$, so müßte nach dem vorigen Satze auch ihr Durchschnitt > 0 sein, was aber nicht der Fall ist. Folglich muß es ein n geben, wofür $A_n = 0$, und dann ist $A = B_n$, $A = \mathfrak{D}(A, S_n)$, also $A \leqq S_n$.

Von diesem Satze gilt eine Art Umkehrung, bei der wir aber statt von einer Folge zunächst von einem beliebigen System oder Komplex von Gebieten sprechen müssen; beide Sätze werden im nächsten Kapitel eine Verschärfung erfahren.

III. (Umkehrung des Borelschen Satzes). **Wenn für jedes System von Gebieten, in deren Summe die Menge A enthalten ist, A bereits in einer Summe von endlich vielen Gebieten dieses Systems enthalten ist, so ist A kompakt und abgeschlossen.**

Nehmen wir erstens an, A sei nicht kompakt, habe also eine divergente Teilmenge B. Zu jedem Punkt a von A existiert dann, weil er nicht Häufungspunkt von B ist, eine Umgebung U_a, die nur endlich viele Punkte von B (vielleicht keinen) enthält. Die Summe aller dieser Umgebungen (Gebiete) U_a schließt A ein, aber eine Summe endlich vieler tut es nicht, da sie nur endlich viele Punkte von B enthält. Das ist ein Widerspruch zu der über A gemachten Voraussetzung, also muß A kompakt sein.

Zweitens sei A nicht abgeschlossen, habe also einen nicht zu A

gehörigen Häufungspunkt x. Für jeden Punkt a von A existiert nach dem Umgebungsaxiom (D) eine Umgebung U_a und eine zugehörige, von a abhängige Umgebung von x, die keinen Punkt mit U_a gemein hat, so daß also x kein α-Punkt von U_a ist. Nach § 3, (8) ist x auch für eine Summe endlich vieler U_a kein α-Punkt; eine solche Summe kann also A nicht einschließen, da x Häufungspunkt von A sein sollte. Die Summe aller U_a aber schließt A ein. Das ist ein Widerspruch zu der über A gemachten Voraussetzung, also ist A abgeschlossen.

Wir übertragen jetzt zwei Hauptbegriffe der elementaren Analysis in die Theorie der Punktmengen, die Begriffe Limes und Konvergenz.

Der Punkt x heißt ein Limes der unendlichen Menge A, wenn jede Umgebung U_x fast alle Punkte (d. h. alle bis auf höchstens endlich viele) der Menge A enthält.

Der Begriff Limes ist also eine Verschärfung des Begriffs Häufungspunkt; ist x ein Limes von A, so ist er gewiß β-Punkt von A. Er ist aber sogar der einzige β-Punkt von A; denn ist $y \neq x$ und wählt man nach dem Axiom (D) U_x und U_y fremd, so enthält U_x fast alle, U_y also nur endlich viele Punkte (eventuell keinen) von A und y ist kein Häufungspunkt von A. Um so mehr ist x der einzige Limes von A. Eine Menge A hat also entweder einen einzigen Limes x oder gar keinen; im ersten Fall nennt man sie konvergent (sie konvergiert nach x) und schreibt

$$x = \lim A.$$

Dies zieht die Gleichung $\{x\} = A_\beta$ nach sich; umgekehrt braucht eine Menge mit einem einzigen Häufungspunkt aber nicht konvergent zu sein (vgl. Satz IV), z. B. auf der geraden Linie die Menge der Punkte $1, 2, \frac{1}{2}, 3, \frac{1}{3}, \ldots$.

Nach der Definition ist $x = \lim A$ auch Limes jeder unendlichen Teilmenge A' von A, also Punkt von A'_β oder A'_α. Umgekehrt ist ein Punkt, der α-Punkt (oder β-Punkt) jeder unendlichen Teilmenge A' der unendlichen Menge A ist, Limes von A; denn gäbe es ein U_x, das nicht fast alle Punkte von A einschließt, so wäre $A' = A - \mathfrak{D}(A, U_x)$ unendlich und hätte x nicht zum α-Punkt. Auch der Limesbegriff läßt sich also auf den Grundbegriff des α-Punktes zurückführen (vgl. § 3 am Ende).

Nach dem soeben Bemerkten ist eine konvergente Menge kompakt, denn jede unendliche Teilmenge hat einen Häufungspunkt. Umgekehrt ist eine kompakte Menge mit einem einzigen Häufungspunkt konvergent. Denn ist A kompakt und x ihr einziger Häufungspunkt, U_x eine Umgebung, so kann $A' = \mathfrak{D}(A, E - U_x)$ nur endlich sein, denn andernfalls hätte A' einen Häufungspunkt y, der von x

verschieden sein muß, weil x nicht einmal α-Punkt von A' ist. U_x enthält also fast alle Punkte von A; $x = \lim A$. Wir haben also die Identität konvergenter Mengen und kompakter Mengen mit einem einzigen Häufungspunkt oder den folgenden Satz bewiesen:

IV. Eine konvergente Menge A ist kompakt und hat einen einzigen Häufungspunkt $x = \lim A$. Eine kompakte Menge A mit einem einzigen Häufungspunkt x ist konvergent und es ist $x = \lim A$.

§ 5. Punkt- und Mengenfolgen.

Von einer abzählbaren Punktmenge zu unterscheiden ist eine Punktfolge

$$\mathfrak{A} = (a_1, a_2, \ldots, a_n, \ldots),$$

d. h. ein Punktkomplex mit der Menge der natürlichen Zahlen als Argument. Zwei Punkte mit verschiedenen Indices brauchen nicht verschieden zu sein; die Menge A der verschiedenen Punkte der Folge $\mathfrak{A}$ ist endlich oder abzählbar.

Die bisher entwickelten Begriffe gestatten eine unmittelbare Übertragung von Mengen auf Folgen. Wir nennen x einen α-Punkt von $\mathfrak{A}$, wenn in jeder Umgebung U_x mindestens ein Punkt der Folge liegt; einen β-Punkt oder Häufungspunkt, wenn in jeder Umgebung U_x unendlich viele (d. h. zu unendlich vielen Indices gehörige) Punkte der Folge liegen.[1] Die Mengen der α- und β-Punkte seien $\mathfrak{A}_\alpha$, $\mathfrak{A}_\beta$; offenbar ist

$$\mathfrak{A}_\alpha = A_\alpha, \quad \mathfrak{A}_\beta \geqq A_\beta,$$

denn ein Häufungspunkt der Menge ist auch Häufungspunkt der Folge, aber nicht umgekehrt. Die aus lauter gleichen Punkten bestehende Folge $(a, a, a, \ldots)$ hat den Punkt a zum Häufungspunkt; die zugehörige Menge $\{a\}$ hat keinen.

x heißt ein Limes der Folge $\mathfrak{A}$, wenn in jeder Umgebung U_x fast alle Punkte der Folge (alle Punkte einer Restfolge $a_n, a_{n+1}, a_{n+2}, \ldots$) liegen. Eine Folge hat höchstens einen einzigen Limes; wenn sie einen hat, so heißt sie konvergent und man schreibt wieder $x = \lim \mathfrak{A}$, in diesem Falle auch $x = \lim a_n$ oder, wenn ein Mißverständnis zu befürchten ist, genauer $x = \lim\limits_n a_n$. Die Folge $(a, a, a, \ldots)$ konvergiert nach a. Jede Teilfolge einer konvergenten Folge konvergiert nach demselben Limes. Ist die zu einer konvergenten Folge $\mathfrak{A}$ gehörige

[1] Wenn wir Punktkomplexe mit beliebigem Argument in derselben Weise behandeln wollten, so wären auch γ-Punkte in Betracht zu ziehen, die es bei Folgen nicht gibt.

Menge A unendlich, so ist A eine konvergente Menge und $\lim A = \lim \mathfrak{A}$. Schreibt man umgekehrt eine konvergente abzählbare[1] Menge A in Gestalt einer Folge $\mathfrak{A}$, indem man ihre Punkte umkehrbar eindeutig den natürlichen Zahlen zuordnet, so ist $\mathfrak{A}$ eine konvergente Folge und $\lim \mathfrak{A} = \lim A$. Eine Folge ohne Häufungspunkt heißt divergent und eine Folge ohne divergente Teilfolge kompakt. Kompakte Folgen mit einem einzigen Häufungspunkt sind konvergent und vice versa. Eine Menge ist dann und nur dann kompakt, wenn alle aus ihren Punkten gebildeten Folgen kompakt sind oder mindestens einen Häufungspunkt haben.

Um eine Anwendung der letzten Bemerkung zu geben, zeigen wir, daß in einem metrischen Raume mit A auch A_α kompakt ist. Wählen wir eine beliebige Folge von Punkten x_n aus A_α und bestimmen zu jedem x_n einen Punkt a_n aus A mit einer Entfernung $\overline{a_n x_n} < \frac{1}{n}$. Da A kompakt ist, hat die Folge der a_n mindestens einen Häufungspunkt x; bei vorgegebenem positivem ϱ ist also für unendlich viele Indices $\overline{x a_n} < \varrho$, $\overline{x x_n} < \varrho + \frac{1}{n}$, also immer noch für unendlich viele $\overline{x x_n} < 2\varrho$, die Folge der x_n hat einen Häufungspunkt x.

In einem metrischen Raume bedeutet die Relation $x = \lim a_n$, daß für jede positive Zahl ϱ fast alle Entfernungen $\overline{x a_n}$ kleiner als ϱ sind, oder also, daß im gewöhnlichen Sinne der Konvergenz von Zahlenfolgen

$$\lim \overline{x a_n} = 0$$

ist. Ferner folgt für einen weiteren Punkt y aus dem Dreiecksaxiom

$$\overline{xy} - \overline{x a_n} \leqq \overline{y a_n} \leqq \overline{xy} + \overline{x a_n}:$$

$$\lim \overline{y a_n} = \overline{yx},$$

und ebenso für zwei konvergente Folgen $\lim a_n = x$, $\lim b_n = y$

$$\lim \overline{a_n b_n} = \overline{xy}.$$

Betrachten wir jetzt eine Mengenfolge $A_1, A_2, \ldots, A_n, \ldots$ und erinnern zunächst an die Begriffe in Kap. I, § 9, wo nur die Angehörigkeit von Elementen zu Mengen in Betracht kam. Wir hatten dort als unteren Limes

$$\text{(I)} \qquad L = \operatorname{Lim\,inf} A_n$$

die Menge der Punkte definiert, die fast allen A_n angehören; als oberen Limes

$$\text{(II)} \qquad M = \operatorname{Lim\,sup} A_n$$

---

[1] Auf dem bisherigen Standpunkt folgt noch nicht, daß eine konvergente Menge stets abzählbar ist (vgl. Kap. VIII, § 2, I).

die Menge der Punkte, die unendlich vielen A_n angehören. Auch die Darstellung dieser Mengen müssen wir hier noch einmal in etwas veränderter Gestalt wiederholen.

Es bedeute

$P = \{p_1, p_2, \dots\}$ eine Menge von fast allen natürlichen Zahlen,

$Q = \{q_1, q_2, \dots\}$ eine Menge von unendlich vielen natürlichen Zahlen,

und wir setzen

$$S_P = \mathfrak{S}_n^P A_n, \quad D_P = \mathfrak{D}_n^P A_n,$$

$$S_Q = \mathfrak{S}_n^Q A_n, \quad D_Q = \mathfrak{D}_n^Q A_n.$$

Dann folgt aus der Definition

(1) $$L = \mathfrak{S} D_P, \quad M = \mathfrak{S} D_Q$$

(die Summe über alle P resp. Q erstreckt). Zwei weitere Formeln gewinnen wir durch Dualität, d. h. Übergang zu den Komplementen. L ist die Menge der Punkte, die nur endlich vielen Komplementen $E - A_n$ angehören, also $E - L$ die Menge derer, die unendlich vielen $E - A_n$ angehören, oder

$$E - \text{Lim inf}\, A_n = \text{Lim sup}\,(E - A_n),$$

ebenso

$$E - \text{Lim sup}\, A_n = \text{Lim inf}\,(E - A_n).$$

Stellt man diese Mengen nach (1) dar und nimmt wieder die Komplemente, so folgt

(2) $$L = \mathfrak{D} S_Q, \quad M = \mathfrak{D} S_P$$

(der Durchschnitt über alle Q, resp. P erstreckt). Damals haben wir nur die erste Formel (1) und die zweite Formel (2) entwickelt, weil diese nur Summe und Durchschnitt abzählbar vieler Mengen zu bilden verlangen, wobei es überdies genügt, statt der sämtlichen P nur die Restfolgen n, $n+1$, $n+2$, ... in Betracht zu ziehen.

Auf unserem jetzigen topologischen Standpunkt ergibt sich nun ein weiter Spielraum für Definitionen neuer Mengen auf Grund der gegebenen Mengenfolge. So könnte man z. B. die Punkte betrachten, die α-Punkte von fast allen Mengen sind; ihre Menge ist $\text{Lim inf}\, A_{n\alpha}$ und nicht identisch mit der Menge der α-Punkte von $\text{Lim inf}\, A_n$, also mit L_α. Ähnliche Begriffsbildungen kann man mit Lim sup und mit β-Punkten, γ-Punkten, inneren Punkten usw. wiederholen. Unter allen diesen Mengen scheinen folgende vier von besonderer Bedeutung zu sein.

x sei ein Punkt, von dem jede Umgebung U_x mit fast allen Mengen A_n Punkte gemein hat. Die Menge dieser Punkte x

(III) $$\overline{L} = \overline{\mathrm{Lim}} \inf A_n$$

heiße der untere abgeschlossene Limes der Mengenfolge.

x sei ein Punkt, von dem jede Umgebung U_x mit unendlich vielen Mengen A_n Punkte gemein hat. Die Menge dieser Punkte x

(IV) $$\overline{M} = \overline{\mathrm{Lim}} \sup A_n$$

heiße der obere abgeschlossene Limes der Mengenfolge.

x sei ein Punkt, von dem eine Umgebung U_x existiert, die fast allen Mengen A_n angehört. Die Menge dieser Punkte x

(V) $$\underline{L} = \underline{\mathrm{Lim}} \inf A_n$$

heiße das untere Limesgebiet der Mengenfolge.

x sei ein Punkt, von dem eine Umgebung U_x existiert, die unendlich vielen Mengen A_n angehört. Die Menge dieser Punkte x

(VI) $$\underline{M} = \underline{\mathrm{Lim}} \sup A_n$$

heiße das obere Limesgebiet der Mengenfolge.

Daß die Mengen (III) und (IV) in der Tat abgeschlossen, die Mengen (V) und (VI) Gebiete sind, geht sowohl aus einer einfachen Überlegung wie aus der nachher folgenden Darstellung hervor.

Wir haben hier also im ganzen sechs Limesbildungen. Aus den Definitionen ergibt sich unmittelbar, daß

$$L \leqq M, \quad \overline{L} \leqq \overline{M}, \quad \underline{L} \leqq \underline{M}.$$

Im Falle der Gleichheit zwischen dem betreffenden unteren und oberen Limes bezeichnen wir diese Menge als Limes schlechthin, also

$$L = M = \mathrm{Lim}\, A_n,$$
$$\overline{L} = \overline{M} = \overline{\mathrm{Lim}}\, A_n,$$
$$\underline{L} = \underline{M} = \underline{\mathrm{Lim}}\, A_n$$

resp. als Limes, abgeschlossenen Limes und Limesgebiet der Mengenfolge.

Ferner ist, ebenfalls in unmittelbarer Folge der Definition,

$$\underline{L} \leqq L \leqq \overline{L}, \quad \underline{M} \leqq M \leqq \overline{M};$$

das Limesgebiet ist in beiden Fällen die kleinste, der abgeschlossene Limes die größte Menge und zwischen beiden liegt der Limes. Mit Rücksicht darauf, daß die unterstrichenen Mengen Gebiete, die überstrichenen abgeschlossen sind, folgt genauer

$$\underline{L} \leqq L_i \leqq L \leqq L_\alpha \leqq \overline{L},$$
$$\underline{M} \leqq M_i \leqq M \leqq M_\alpha \leqq \overline{M},$$

ohne daß etwa $\underline{L} = L_i$ usw. zu sein brauchte.

Eine Änderung, Weglassung oder Hinzufügung endlich vieler Mengen ist auf die Limesmengen ohne Einfluß. Betrachten wir von der Mengenfolge eine Teilfolge $A_{n_1}, A_{n_2}, \ldots$ und bezeichnen deren sechs Limites durch Anhängung des Index 1, so folgt unmittelbar

$$L \leqq L_1 \leqq M_1 \leqq M,$$
$$\overline{L} \leqq \overline{L_1} \leqq \overline{M_1} \leqq \overline{M},$$
$$\underline{L} \leqq \underline{L_1} \leqq \underline{M_1} \leqq \underline{M},$$

woraus hervorgeht, daß die Existenz von Lim, $\overline{\text{Lim}}$ oder $\underline{\text{Lim}}$ für die ganze Folge die Existenz desselben Limes für die Teilfolge nach sich zieht.

Die Darstellung der vier neu hinzugekommenen Limites beginnen wir mit den Limesgebieten. Ein Punkt x der Menge $\underline{L}$ ist offenbar innerer Punkt eines der Durchschnitte D_P aus fast allen A_n [1], also Punkt von D_{Pi}, und umgekehrt; ebenso ist $\underline{M}$ die Menge der inneren Punkte der Durchschnitte D_Q aus unendlich vielen A_n. Demnach ist

(3) $$\underline{L} = \mathfrak{S} D_{Pi}, \quad \underline{M} = \mathfrak{S} D_{Qi}.$$

Die beiden andern Mengen gewinnen wir wieder durch Komplementbildung. Ist x ein Punkt von $\overline{L}$, so gehört jede Umgebung U_x nur endlich vielen Komplementen $E - A_n$ an; ein Punkt y von $E - \overline{L}$ hat also eine Umgebung U_y, die unendlich vielen $E - A_n$ angehört, d. h. $E - \overline{L}$ ist das obere Limesgebiet der Mengen $E - A_n$. Für einen Punkt x von $\overline{M}$ hat jede Umgebung U_x Punkte mit unendlich vielen A_n gemein; für einen Punkt y von $E - \overline{M}$ gibt es also eine Umgebung U_y, die nur mit endlich vielen A_n Punkte gemein hat, also fast allen $E - A_n$ angehört: $E - \overline{M}$ ist das untere Limesgebiet der Mengen $E - A_n$. Also ist

$$E - \overline{\text{Lim}} \inf A_n = \underline{\text{Lim}} \sup (E - A_n),$$
$$E - \overline{\text{Lim}} \sup A_n = \underline{\text{Lim}} \inf (E - A_n).$$

Stellen wir die Mengen rechterhand nach (3) dar und nehmen die Komplemente (wobei nach § 3, außer den sonstigen Vertauschungen, die Menge der inneren Punkte durch die Menge der α-Punkte zu ersetzen ist), so ergibt sich

(4) $$\overline{L} = \mathfrak{D} S_{Q\alpha}, \quad \overline{M} = \mathfrak{D} S_{P\alpha}.$$

Aus diesen Darstellungen folgt ein Teil des bereits Gesagten noch einmal, z. B. daß die Mengen (3) als Summen von Gebieten wieder Gebiete, die Mengen (4) als Durchschnitte abgeschlossener Mengen

---

[1] Dazu genügt nicht etwa, daß er innerer Punkt fast aller A_n sei.

wieder abgeschlossen sind usw. Die nach P zu erstreckenden Summen und Durchschnitte (bei denen man sich wieder auf Restfolgen beschränken kann) umfassen nur abzählbar viele Mengen, die nach Q hingegen unabzählbar viele; die Mengen $\underline{L}$ und $\overline{M}$ sind also in dieser Hinsicht einfacher als die beiden andern.

Ebenfalls aus den Formeln (3), (4) oder direkt aus der Definition folgt, daß die Mengen A_n bei der Bildung der Limesgebiete durch die größten darin enthaltenen Gebiete A_{ni}, bei der Bildung der abgeschlossenen Limites durch die kleinsten sie enthaltenden abgeschlossenen Mengen $A_{n\alpha}$ ersetzt werden dürfen. Denn für

$$D = \mathfrak{D}(A_1, A_2, \ldots), \quad \varDelta = \mathfrak{D}(A_{1i}, A_{2i}, \ldots)$$

ist $D_i \leqq \varDelta \leqq D$, folglich $D_i = \varDelta_i$, und in (3) bleiben also D_{Pi}, D_{Qi} bei Ersetzung der A_n durch die A_{ni} unverändert; ferner für

$$S = \mathfrak{S}(A_1, A_2, \ldots), \quad \Sigma = \mathfrak{S}(A_{1\alpha}, A_{2\alpha}, \ldots)$$

ist $S_\alpha \geqq \Sigma \geqq S$ und $S_\alpha = \Sigma_\alpha$, woraus das Entsprechende für (4) folgt. Bei der Bildung der Limesgebiete kann jede Menge A_n durch eine Menge mit denselben inneren Punkten, bei der Bildung der abgeschlossenen Limites durch eine Menge mit denselben α-Punkten ersetzt werden.

Beispiele.

(α) Für die Folge $A, B, A, B, \ldots$ ist stets $S_P = \mathfrak{S}(A, B)$, $D_P = \mathfrak{D}(A, B)$, während $S_Q = \mathfrak{S}(A, B)$ oder A oder B und $D_Q = \mathfrak{D}(A, B)$ oder A oder B sein kann. Nach (1) ist

$$L = \mathfrak{D}(A, B), \quad M = \mathfrak{S}\big(\mathfrak{D}(A, B), A, B\big) = \mathfrak{S}(A, B)$$

und nach (2)

$$L = \mathfrak{D}\big(\mathfrak{S}(A, B), A, B\big) = \mathfrak{D}(A, B), \quad M = \mathfrak{S}(A, B);$$

nach (3)

$$\underline{L} = \mathfrak{D}(A, B)_i = \mathfrak{D}(A_i, B_i),$$
$$\underline{M} = \mathfrak{S}\big(\mathfrak{D}(A, B)_i, A_i, B_i\big) = \mathfrak{S}(A_i, B_i)$$

und nach (4)

$$\overline{L} = \mathfrak{D}\big(\mathfrak{S}(A, B)_\alpha, A_\alpha, B_\alpha\big) = \mathfrak{D}(A_\alpha, B_\alpha),$$
$$\overline{M} = \mathfrak{S}(A, B)_\alpha = \mathfrak{S}(A_\alpha, B_\alpha).$$

Also

$$L = \mathfrak{D}(A, B), \quad M = \mathfrak{S}(A, B),$$
$$\underline{L} = \mathfrak{D}(A_i, B_i), \quad \underline{M} = \mathfrak{S}(A_i, B_i),$$
$$\overline{L} = \mathfrak{D}(A_\alpha, B_\alpha), \quad \overline{M} = \mathfrak{S}(A_\alpha, B_\alpha).$$

Hier ist zwar $\underline{L} = L_i$, $\overline{M} = M_\alpha$, aber im allgemeinen nur $\underline{M} \leqq M_i$, $\overline{L} \geqq L_\alpha$.

Für $A = B$, wenn also alle Mengen identisch sind, ist

$$L = M = A, \quad \underline{L} = \underline{M} = A_i, \quad \overline{L} = \overline{M} = A_\alpha.$$

Wir haben hier die Darstellung nach den Formeln (1)—(4) ausführlich gegeben; bei den folgenden Beispielen überlassen wir sie, oder die Herleitung der Limesmengen aus der Definition, dem Leser.

(β) Die A_n mögen eine absteigende Folge bilden: $A_1 \geqq A_2 \geqq \cdots$

$$L = M = \mathfrak{D}(A_1, A_2, \ldots) = D,$$

$$\underline{L} = \underline{M} = D_i,$$

$$\overline{L} = \overline{M} = \mathfrak{D}(A_{1\alpha}, A_{2\alpha}, \ldots) \geqq D_\alpha.$$

(γ) Die A_n bilden eine aufsteigende Folge: $A_1 \leqq A_2 \leqq \cdots$

$$L = M = \mathfrak{S}(A_1, A_2, \ldots) = S,$$

$$\underline{L} = \underline{M} = \mathfrak{S}(A_{1i}, A_{2i}, \ldots) \leqq S_i,$$

$$\overline{L} = \overline{M} = S_\alpha.$$

(δ) Wenn die A_n paarweise fremd sind, ist jedenfalls

$$L = M = \underline{L} = \underline{M} = 0.$$

(ε) $A = \{a_1, a_2, \ldots\}$ sei eine abzählbare Menge und $A_n = \{a_n\}$ bestehe aus dem einen Punkt a_n. Dann ist $\overline{M} = A_\beta$, hingegen $\overline{L} = \lim A$ oder $= 0$, je nachdem die Menge A konvergent ist oder nicht. Der abgeschlossene Limes $\overline{L} = \overline{M}$ existiert dann und nur dann, wenn die Menge A konvergent oder divergent ist.

Setzt man dagegen $A_n = \{a_1, a_2, \ldots, a_n\}$, so ist $\overline{L} = \overline{M} = A_\alpha$. Jede abgeschlossene Menge, in der eine abzählbare Menge dicht ist (§ 8), kann also als abgeschlossener Limes endlicher Mengen dargestellt werden.

Setzt man drittens $A_n = \{a_n, a_{n+1}, \ldots\}$, so ist $\overline{L} = \overline{M} = A_\beta$.

Diesen Beispielen allgemeiner Art fügen wir noch einige spezielle hinzu.

(ζ) In der euklidischen Ebene sei A_n die Gerade $x = \frac{1}{n}$; dann ist $\overline{L} = \overline{M}$ die Ordinatenachse $x = 0$. Für die Folge der Geraden $x = 1, 2, \frac{1}{2}, 3, \frac{1}{3}, \ldots$ ist $\overline{L} = 0$, $\overline{M}$ die Ordinatenachse.

(η) A_n sei die gebrochene Linie $a c_n b$, die die Punkte mit den rechtwinkligen Koordinaten $(-1, 0)$ $(0, n)$ $(1, 0)$ verbindet. $\overline{L} = \overline{M}$ besteht aus den beiden Halbgeraden $x = \pm 1$, $y \geqq 0$.

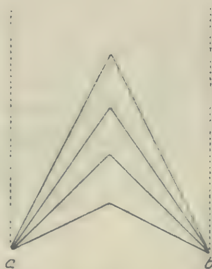


Fig. 4.

§ 6. Relativbegriffe.

Die Menge A heiße in M abgeschlossen (relativ abgeschlossen), wenn

(1) $$A = \mathfrak{D}(M, F)$$

der Durchschnitt von M mit einer abgeschlossenen Menge F ist. Dazu ist die Bedingung

(2) $$A = \mathfrak{D}(M, A_\alpha)$$

notwendig und hinreichend. Denn aus (1) folgt

$$A \leqq F, \quad A_\alpha \leqq F_\alpha = F,$$

$$A = \mathfrak{D}(A, A_\alpha) = \mathfrak{D}(M, F, A_\alpha) = \mathfrak{D}(M, A_\alpha),$$

und aus (2) folgt, daß A der Durchschnitt von M mit einer abgeschlossenen Menge A_α ist.

Die Bedingung (2) ist mit

$$A \geqq \mathfrak{D}(M, A_\beta)$$

gleichwertig, wenn wir A als Teilmenge von M voraussetzen, und besagt: die Teilmenge A von M ist in M abgeschlossen dann und nur dann, wenn A alle seine zu M gehörigen Häufungspunkte enthält, oder wenn $M - A$ keinen Häufungspunkt von A enthält.

Folgende Sätze sind unmittelbar evident:

Die Summe endlich vieler und der Durchschnitt beliebig vieler in M abgeschlossener Mengen ist wieder in M abgeschlossen.

Jede Menge ist in sich selbst abgeschlossen. Ist A in B, B in C abgeschlossen, so auch A in C. Ist $A \leqq B \leqq C$ und A in C abgeschlossen, so auch A in B. Ist A in B abgeschlossen und B (absolut, d. h. in E) abgeschlossen, so ist auch A abgeschlossen. Ist $A \leqq B$ und A abgeschlossen, so ist auch A in B abgeschlossen.

Der Kern jeder beliebigen Menge M ist in M abgeschlossen.

Denn ist $M = M_k + M_s$, M_k der Kern oder die größte insichdichte Teilmenge von M, so kann $M - M_k = M_s$ keinen Häufungspunkt h von M_k enthalten, da sonst $M_k + \{h\}$ insichdicht und $> M_k$ wäre.

Der Kern einer abgeschlossenen Menge ist demnach selbst abgeschlossen und gleichzeitig insichdicht, also:

Der Kern jeder abgeschlossenen Menge ist perfekt.

Die Menge B heiße ein Relativgebiet von M, wenn

(3) $$B = \mathfrak{D}(M, G)$$

der Durchschnitt von M mit einem Gebiet G ist. Dann ist $F = E - G$ abgeschlossen, $A = M - B = \mathfrak{D}(M, F)$ in M abgeschlossen und vice versa, d. h. Relativgebiete von M und in M abgeschlossene

Mengen sind Komplemente voneinander (in M). Über Relativgebiete gelten entsprechende Sätze wie über relativ abgeschlossene Mengen, z. B.:

Der Durchschnitt endlich vieler und die Summe beliebig vieler Relativgebiete von M ist wieder ein Relativgebiet von M.

Für das Relativgebiet (3) gelten außer den allgemeinen Formeln § 3, (7) für Durchschnitte noch die folgenden:

(4) $$B_\alpha \geqq \mathfrak{D}(M_\alpha, G), \quad B_\beta \geqq \mathfrak{D}(M_\beta, G), \quad B_\gamma \geqq \mathfrak{D}(M_\gamma, G).$$

Denn ist $A = \mathfrak{D}(M, F)$ das Komplement von B in M, wobei also

$$M = A + B, \quad E = F + G,$$

G ein Gebiet und F abgeschlossen ist, so ist wegen $F = F_\alpha \geqq F_\beta \geqq F_\gamma$

$$A_\alpha \leqq \mathfrak{D}(M_\alpha, F_\alpha) = \mathfrak{D}(M_\alpha, F)$$

und analog:

$$A_\alpha \leqq \mathfrak{D}(M_\alpha, F), \quad A_\beta \leqq \mathfrak{D}(M_\beta, F), \quad A_\gamma \leqq \mathfrak{D}(M_\gamma, F).$$

Daraus folgen die Formeln (4) nach einer schon mehrfach angewandten Schlußweise, nämlich: ein Punkt von $\mathfrak{D}(M_\alpha, G)$ ist Punkt von $M_\alpha = \mathfrak{S}(A_\alpha, B_\alpha)$, aber nicht Punkt von A_α, da er sonst Punkt von $\mathfrak{D}(M_\alpha, F)$ wäre, folglich ist er Punkt von B_α, womit die erste Formel (4) bewiesen ist.

Ein Relativgebiet von einer insichdichten Menge ist wieder insichdicht.[1] Denn aus $M_\beta \geqq M$ folgt nach (4) $B_\beta \geqq B$.

Die Begriffe des Relativgebiets und der relativ abgeschlossenen Menge sind noch einer etwas systematischeren Auffassung fähig, indem wir eine ganze Relativtheorie entwickeln, die sich, statt auf die Menge E, auf eine Teilmenge von ihr als zugrunde liegenden Raum bezieht.

Sobald eine Menge E mit einem System von Teilmengen (Umgebungen) U_x vorliegt, das den vier Umgebungsaxiomen genügt, so sind für eine beliebige Menge $A \leqq E$ die entsprechenden Mengen A_α, A_i usw. definiert. Ist A außerdem Teilmenge einer anderen Menge M, für deren Punkte y ein System von Umgebungen V_y mit den gleichen Axiomen gegeben ist, so erhalten wir ein neues System von Mengen $A_{(\alpha)}$, $A_{(i)}$ usw. und eine neue, mit der ersten durchaus gleichlautende Theorie, in der alle Sätze der ursprünglichen formal richtig bleiben, aber eine ganz andere Bedeutung erlangen können. Z. B. ist nach wie vor die Summe von Gebieten wieder ein Gebiet, aber eine Menge kann im ersten System ein Gebiet sein und im

[1] Im euklidischen Raume sind Gebiete insichdicht; indessen braucht der Durchschnitt zweier insichdichter Mengen im allgemeinen keine insichdichte Menge zu sein.

zweiten nicht, wenn nämlich $A = A_i \dotplus A_{(i)}$. Solche Überlegungen machen sich z. B. notwendig, wenn wir eine ebene Punktmenge gleichzeitig als räumliche auffassen wollen.

Der wichtigste Fall ist, daß, bei gegebener Menge E und bestimmten Umgebungen U_x, M eine Teilmenge von E ist und die neuen Umgebungen als die entsprechenden Teilmengen der alten definiert werden; wir ordnen also jedem Punkte y von M die Umgebungen

$$V_y = \mathfrak{D}(U_y, M) \tag{5}$$

zu, die evidentermaßen die vier Umgebungsaxiome erfüllen. Ist z. B. E ein metrischer Raum, so ist V_y die Umgebung mit gleichem Radius ϱ wie U_y, nämlich die Menge der Punkte von M, deren Entfernung von y kleiner als ϱ ist.

Nachdem hiermit die Umgebungen fixiert sind, hängt alles weitere nur von der Menge M ab, und jeder solchen (nichtverschwindenden) Menge M entspricht eine Relativtheorie, in der M statt E die Rolle des umfassenden Raumes übernommen hat. Ist A Teilmenge von M, so nennen wir also einen Punkt y von M einen relativen α-, β-, γ-Punkt von A, wenn jede Relativumgebung V_y mindestens einen resp. unendlich viele resp. unabzählbar viele Punkte von A enthält, oder wenn immer

$$\mathfrak{D}(V_y, A) = \mathfrak{D}(U_y, M, A) = \mathfrak{D}(U_y, A)$$

von Null verschieden resp. unendlich resp. unabzählbar ist. Das besagt, daß y ein (zu M gehöriger) absoluter α-, β-, γ-Punkt von A ist, und für die Mengen dieser Punkte ergibt sich also einfach

$$A_{(\alpha)} = \mathfrak{D}(A_\alpha, M), \quad A_{(\beta)} = \mathfrak{D}(A_\beta, M), \quad A_{(\gamma)} = \mathfrak{D}(A_\gamma, M). \tag{6}$$

Nennen wir demgemäß A relativ abgeschlossen oder in M abgeschlossen, wenn $A = A_{(\alpha)}$, so stimmt dies mit unserer anfänglichen Definition überein.

Die Menge der zu A gehörigen relativen Häufungspunkte von A ist

$$A_{(h)} = \mathfrak{D}(A, A_{(\beta)}) = \mathfrak{D}(A, M, A_\beta) = \mathfrak{D}(A, A_\beta) = A_h,$$

also von M unabhängig und gleich der Menge der zu A gehörigen absoluten Häufungspunkte von A; man kann auch sagen, daß A_h die relative Ableitung von A bezogen auf A selbst ist. Ebenfalls von dem umfassenden Raume M unabhängig ist demgemäß die Definition insichdichter Mengen ($A = A_h$), folglich auch der Kern A_k und die Mengen A_j und A_s.

Der Punkt y von M heißt ein relativ innerer Punkt von A, wenn eine Umgebung $V_y \leqq A$ existiert. Die Menge $A_{(i)}$ der relativ

inneren Punkte ergibt sich durch die ins Relative übertragene Komplementbildung folgendermaßen. Setzen wir

$$M = A + B, \quad E = A' + B,$$

also

$$A' = A + (E - M),$$

so ist

(7) $$A_{(i)} = M - B_{(\alpha)} = \mathfrak{D}(E - B_\alpha, M) = \mathfrak{D}(A'_i, M).$$

Wegen $A' \geqq A$ ist $A_{(i)} \geqq \mathfrak{D}(A_i, M)$, $A_{(i)}$ also im allgemeinen nicht die Menge der zu M gehörigen absolut inneren Punkte von A, sondern größer als diese; auch absolute Randpunkte können zu relativ inneren Punkten werden.

Nennen wir A ein Relativgebiet von M, wenn $A = A_{(i)}$, also das Komplement $B = B_{(\alpha)}$ in M abgeschlossen ist, so stimmt dies mit der zuvor gegebenen Erklärung überein.

Die Übertragung der weiteren Begriffe ins Relative überlassen wir dem Leser. Z. B. wird man A in M kompakt nennen, wenn jede unendliche Teilmenge von A einen Häufungspunkt in M hat (also kompakt schlechthin = kompakt in E); eine in M abgeschlossene und in M kompakte Menge ist in sich selbst kompakt (über das Umgekehrte vgl. Kap. VIII, § 2).

Ist speziell $M = E_m$ ein euklidischer Raum von m Dimensionen, der in einem euklidischen Raum $E = E_n$ von n Dimensionen $(n > m)$ enthalten ist, und legt man die gewöhnliche Definition von U_x (Inneres einer n-dimensionalen Kugel vom Radius ϱ) zugrunde, so erhält auch V_y nach (5) die gewöhnliche Bedeutung (Inneres einer m-dimensionalen Kugel vom Radius ϱ). Die (absolute) Theorie der Punktmengen in E_m wird dann, wenn wir diese Punktmengen jetzt als Punktmengen in E_n betrachten, zur Relativtheorie bezüglich E_m; was wir auf dem m-dimensionalen Standpunkt eine abgeschlossene Menge oder ein Gebiet nennen, ist auf dem n-dimensionalen eine in E_m abgeschlossene Menge oder ein Relativgebiet von E_m. Als Besonderheit kommt hier noch hinzu, daß E_m (in E_n) abgeschlossen ist, also für $A \leqq E_m$ auch $A_\alpha \leqq E_m$, und daher nach (6)

$$A_{(\alpha)} = A_\alpha, \quad A_{(\beta)} = A_\beta, \quad A_{(\gamma)} = A_\gamma.$$

Diese Mengen hängen also im gegenwärtigen Falle von dem Raume, in dem wir A betrachten, nicht ab; eine in E_m abgeschlossene Menge ist auch absolut (in E_n) abgeschlossen. Nicht so steht es mit $A_{(i)}$, dessen Verschiedenheit von $A_i = 0$ im allgemeinen bestehen bleibt und sich durch die Formel (7) erklärt, und mit der Definition des Gebietes, die also vom umgebenden Raume abhängt; eine von 0 verschiedene Teilmenge A von E_m ist in E_n niemals ein Gebiet, kann aber eins in E_m sein (wenn $E_m - A$ abgeschlossen ist; $E_n - A$ ist nicht abgeschlossen).

§ 7. Zusammenhang.

Wir definieren[1]:

Eine von Null verschiedene Menge A heißt zusammenhängend, wenn sie sich nicht in zwei fremde, von Null verschiedene, in A abgeschlossene Teilmengen spalten läßt.

Um also den Zusammenhang einer Menge A zu beweisen, hat man zu zeigen, daß bei jeder Zerlegung $A = A_1 + A_2$, wo A_1 und A_2 in A abgeschlossen sind, die eine dieser Teilmengen verschwinden muß.

Da A_1 als Komplement von A_2 ein Relativgebiet ist, ebenso A_2 als Komplement von A_1, so kann man auch sagen: eine Menge heißt zusammenhängend, wenn sie sich nicht in zwei (fremde, von Null verschiedene) Relativgebiete spalten läßt.

Insbesondere ist eine abgeschlossene Menge zusammenhängend, wenn sie sich nicht in zwei abgeschlossene Mengen, ein Gebiet, wenn es sich nicht in zwei Gebiete spalten läßt.

Eine Menge, die nur einen Punkt enthält, sehen wir im Einklang mit der Definition als zusammenhängend an.

Wir leiten einige einfache Sätze ab, die teilweise nur hinreichende, nicht notwendige Bedingungen des Zusammenhangs aussprechen. Dabei wird mehrmals von der Bemerkung Gebrauch gemacht, daß eine Zerlegung

$$A = \mathfrak{D}(A, F_1) + \mathfrak{D}(A, F_2) = A_1 + A_2$$

in zwei relativ abgeschlossene Mengen auch für jede Teilmenge B von A eine entsprechende Zerlegung

$$\begin{aligned} B &= \mathfrak{D}(B, A_1) + \mathfrak{D}(B, A_2) \\ &= \mathfrak{D}(B, F_1) + \mathfrak{D}(B, F_2) = B_1 + B_2 \end{aligned}$$

in zwei relativ abgeschlossene Mengen hervorruft.

I. Wenn je zwei Punkte von A einer zusammenhängenden Teilmenge von A angehören, so ist A selbst zusammenhängend.

Angenommen, A sei nicht zusammenhängend, also $A = A_1 + A_2$ eine Zerlegung in zwei nichtverschwindende, relativ abgeschlossene Summanden. Sind dann x_1, x_2 zwei beliebige Punkte von A_1, A_2, so gehören sie einer zusammenhängenden Teilmenge B von A an; andererseits entspricht jener Zerlegung von A eine Zerlegung $B = B_1 + B_2$, wo B_1 und B_2 in B abgeschlossen und von Null verschieden sind (weil sie die Punkte x_1, x_2 enthalten). Dies wider-

[1] Für den Zusammenhang sind verschiedene Definitionen üblich; die im Text gegebene deckt sich mit keiner von ihnen, scheint uns aber die natürlichste und allgemeinste zu sein.

spricht aber dem Zusammenhange von B; also ist die Annahme, A sei unzusammenhängend, zu verwerfen.

Wir werden später sehen, daß z. B. eine geradlinige Strecke zusammenhängend ist; danach ist jede konvexe Menge A, d. h. eine solche, in der je zwei Punkte durch eine zu A gehörige Strecke verbunden werden können, zusammenhängend, z. B. die ganze Ebene, das Innere eines Kreises u. dgl.

II. Die Summe zweier zusammenhängender Mengen, die mindestens einen Punkt gemeinsam haben, ist selbst zusammenhängend.

Wir können die Mengen in der Form $A+D$, $B+D$ annehmen, wo D ihr von Null verschiedener Durchschnitt und A, B, D paarweise fremd sind; ihre Summe ist dann

$$S = A + B + D.$$

Sei $S = S_1 + S_2$ eine Zerlegung in zwei relativ abgeschlossene Mengen, und $A = A_1 + A_2$, $B = B_1 + B_2$, $D = D_1 + D_2$ die entsprechenden Zerlegungen. Dann sind $A_1 + D_1$ und $A_2 + D_2$ in $A + D$ abgeschlossen, und weil $A + D = (A_1 + D_1) + (A_2 + D_2)$ zusammenhängend vorausgesetzt war, so muß einer dieser Summanden Null sein, etwa $A_1 + D_1 = 0$, also $A_1 = D_1 = 0$. Wegen $D > 0$ ist dann $D_2 > 0$, und da ebenso eine der Mengen $B_1 + D_1$, $B_2 + D_2$ verschwinden muß, so muß $B_1 + D_1 = 0$, $B_1 = D_1 = 0$ sein. Damit folgt $S_1 = 0$, die Zerlegung von S in zwei relativ abgeschlossene Mengen bedingt also das Verschwinden der einen: S ist zusammenhängend.

III. Die Summe beliebig vieler zusammenhängender Mengen, die paarweise mindestens einen Punkt gemein haben, ist wieder zusammenhängend.

Es sei $S = \mathfrak{S}(A, B, C, \ldots)$. Zwei Punkte von S gehören entweder einem Summanden A an, der eine zusammenhängende Menge ist, oder sie gehören zwei Summanden A, B an, deren Summe $\mathfrak{S}(A, B)$ nach II ebenfalls eine zusammenhängende Menge ist. In jedem Falle gehören sie also einer zusammenhängenden Teilmenge von S an, und nach I ist S selbst zusammenhängend.

Ist insbesondere $D = \mathfrak{D}(A, B, C, \ldots)$, der Durchschnitt aller der zusammenhängenden Mengen $A, B, C, \ldots$, von Null verschieden, so ist S zusammenhängend.

Sei A eine beliebige von 0 verschiedene Menge. Unter einer Komponente von A verstehen wir eine größte zusammenhängende Teilmenge von A, d. h. eine solche, die in keiner anderen enthalten ist. Die Existenz solcher ist auf Grund von III leicht zu folgern. Ist nämlich p ein Punkt von A, so gibt es jedenfalls zusammenhängende Teilmengen, die p enthalten (wenn

keine andere, so wenigstens die Menge $\{p\}$, die aus p allein besteht). Die Summe aller zusammenhängenden Teilmengen, die p enthalten, sei $P = A(p)$; dann ist P eine Komponente, denn nach III ist P selbst zusammenhängend, und jede zusammenhängende Menge $\geqq P$ enthält p, ist also ein Summand der Summe P und $\leqq P$, folglich $= P$.

Wenn zwei Komponenten P, Q einen Punkt gemein haben, so sind sie identisch. Denn nach II ist dann auch die Summe $\mathfrak{S}(P, Q)$ zusammenhängend und $\geqq P$; sie kann nicht $> P$ sein, weil P sonst keine größte zusammenhängende Menge wäre, und demnach ist $\mathfrak{S}(P, Q) = P = Q$. Ein Punkt von A kann also höchstens einer Komponente angehören, nach dem vorhin Gesagten gehört er aber auch wirklich einer an. Danach zerfällt A in Komponenten derart, daß, wenn man aus jeder Komponente einen Punkt p auswählt und diese Punkte die Teilmenge $B = \{p, q, \ldots\}$ von A bilden,

$$A = \sum_p^B A(p) = A(p) + A(q) + \ldots = P + Q + \ldots$$

ist. Die Mächtigkeit der Menge B heiße die **Komponentenzahl** von A; sie ist für eine zusammenhängende Menge 1, für eine unzusammenhängende Menge > 1, wobei sie endlich oder ein Alef sein kann. Man beachte wohl, daß die Komponenten auch aus einzelnen Punkten bestehen können; sogar wissen wir bis jetzt nicht, ob es zusammenhängende Mengen mit mehr als einem Punkt wirklich gibt.

IV. **Ist A zusammenhängend, so ist auch die Menge A_α und jede Menge zwischen A und A_α zusammenhängend.**

Mit andern Worten: aus einer zusammenhängenden Menge entsteht durch Hinzufügung beliebig vieler Häufungspunkte wieder eine zusammenhängende Menge.

Sei $A \leqq B \leqq A_\alpha$ und $B = A + H$, so daß H aus Häufungspunkten von A besteht, die nicht zu A gehören. Einer Zerlegung $B = B_1 + B_2$ in zwei relativ abgeschlossene Mengen entsprechen die Zerlegungen $A = A_1 + A_2$, $H = H_1 + H_2$. Wegen des Zusammenhanges der Menge A muß einer ihrer Summanden Null sein, etwa $A_1 = 0$, $A_2 = A$,

$$B = H_1 + (A + H_2).$$

Dann muß aber auch $H_1 = 0$ sein, denn für $H_1 > 0$ wäre $A + H_2$ nicht in B abgeschlossen, weil ein Teil ihrer Häufungspunkte nicht zu ihr selbst, wohl aber zu B gehören würde. Folglich ist $B_1 = 0$, B zusammenhängend.

V. **Die Komponenten einer Menge A sind in A abgeschlossen.**

Denn ist P eine Komponente von A, so kann $A - P$ keinen Häufungs-

punkt h von P enthalten, da sonst $P + \{h\}$ nach IV zusammenhängend, P also keine größte zusammenhängende Teilmenge wäre.[1]

VI. **Eine zusammenhängende Menge, die mehr als einen Punkt enthält, ist insichdicht.**

Gleichbedeutend damit ist die Behauptung: wenn A einen isolierten Punkt a und außerdem noch andere Punkte enthält, so ist A unzusammenhängend. Ist U_a eine Umgebung, die keinen Punkt von A außer a enthält, so ist

$$A = \mathfrak{D}(A, U_a) + \mathfrak{D}(A, E - U_a) = \{a\} + A_1 .$$

Hier ist A_1 in A abgeschlossen, weil $E - U_a$ abgeschlossen ist (als Komplement des Gebietes U_a); $\{a\}$ ist als endliche Menge abgeschlossen, also auch in A abgeschlossen. Beide Mengen sind > 0, also ist A unzusammenhängend. Es geht hieraus zugleich hervor, daß $\{a\}$ für sich allein eine Komponente von A ist.

Eine **abgeschlossene zusammenhängende** Menge, die aus mehr als einem Punkt besteht, ist nach VI **perfekt**. Die Komponenten einer abgeschlossenen Menge sind nach V abgeschlossen, also entweder Mengen mit einem einzigen Punkt oder perfekt.

VII. **Enthält eine zusammenhängende Menge Punkte von Komplementärmengen A, B ($E = A + B$), so enthält sie auch einen Punkt der Grenze $A_g = B_g$.**

Ist C diese zusammenhängende Menge, so ist

$$C = \mathfrak{D}(C, A) + \mathfrak{D}(C, B)$$

und beide Summanden sind von Null verschieden. Wäre nun

$$\mathfrak{D}(C, A_g) = \mathfrak{D}(C, A_r) + \mathfrak{D}(C, B_r) = 0,$$

so wären $\mathfrak{D}(C, A) = \mathfrak{D}(C, A_i)$ und $\mathfrak{D}(C, B) = \mathfrak{D}(C, B_i)$ Relativgebiete von C, und C also nicht zusammenhängend.

In der euklidischen Ebene ist jede zusammenhängende Menge, die mehr als einen Punkt enthält, von der Mächtigkeit des Kontinuums. Denn wenn a und b zwei verschiedene Punkte von C sind, so ziehe man eine Gerade G derart, daß a und b auf verschiedenen Seiten von G liegen; A sei die Halbebene (inklusive G), in der a liegt, B die Halbebene (exklusive G), in der b liegt. Dann ist $A_r = G$, $B_r = 0$, $A_g = B_g = G$, und C muß also G in mindestens einem Punkte treffen. Da man eine mit dem Kontinuum äquivalente Menge paralleler Geraden dieser Art ziehen kann, so enthält C (mindestens und auch höchstens) $\aleph = 2^{\aleph_0}$ Punkte.

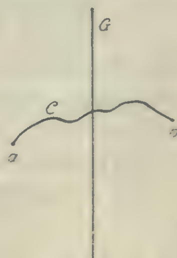


Fig. 5.

Danach ist jede endliche oder abzählbare Menge, die Mengen aus einem

[1] Dagegen braucht $A - P$ in A nicht abgeschlossen zu sein.

Punkt ausgenommen, gewiß unzusammenhängend, und die Komponenten einer solchen Menge bestehen aus ihren einzelnen Punkten. Man sieht, wie der Beweis von der Ebene auf den n-dimensionalen Raum E_n zu übertragen ist; auf der geraden Linie E_1 enthält eine zusammenhängende Menge zugleich mit zwei Punkten auch jeden zwischenliegenden Punkt.

In einem metrischen Raum enthält eine zusammenhängende Menge zugleich mit zwei verschiedenen Punkten a, b für jede positive Zahl $\varrho < \overline{ab}$ mindestens einen Punkt x mit $\overline{ax} = \varrho$, ist also mindestens von der Mächtigkeit des Kontinuums. Denn andernfalls zerfiele sie in zwei Relativgebiete, nämlich die Menge ihrer Punkte mit $\overline{ax} < \varrho$ und mit $\overline{ax} > \varrho$.

Es sei A eine nichtverschwindende Menge und

$$A = A_1 + A_2 = B_1 + B_2 = C_1 + C_2 = \ldots$$

möge alle Zerlegungen von A in zwei fremde, relativ abgeschlossene Mengen bedeuten (auch $A + 0$ mitgerechnet). Ein bestimmter Punkt p gehöre zu $A_1, B_1, C_1, \ldots$. Die Menge

$$A_p = \mathfrak{D}(A_1, B_1, C_1, \ldots),$$

die wieder in A abgeschlossen ist, ist die Menge aller Punkte, die bei jeder Zerlegung demselben Summanden wie p angehören; wir nennen sie die zu p gehörige Quasikomponente von A.[1] Es ist unmittelbar ersichtlich, daß, wenn p und p' stets demselben Summanden angehören, ebenso p' und p'', auch p und p'' es tun, daß folglich jeder Punkt von A einer und nur einer Quasikomponente angehört und damit eine bestimmte Zerlegung

$$A = A_p + A_q + \cdots$$

in Quasikomponenten hervorgerufen wird. Eine zusammenhängende Teilmenge von A gehört bei jeder Zerlegung einem Summanden an, da sie sonst in zwei relativ abgeschlossene Teilmengen zerspalten würde; insbesondere ist die zu p gehörige Komponente $A(p)$ in der zu p gehörigen Quasikomponente A_p enthalten: $A(p) \leqq A_p$. Wohl aber können mehrere Komponenten sich zu einer Quasikomponente vereinigen, wie ein Beispiel zeigen wird; die Anzahl der Quasikomponenten ist höchstens gleich der Anzahl der Komponenten. Ist das Komplement $A - A_p$ einer Quasikomponente in A abgeschlossen, so ist A_p zusammenhängend, also eine Komponente; denn wäre $A_p = B + C$ eine Zerlegung in zwei nichtverschwindende Summanden, die in A_p, also auch in A abgeschlossen sind, so wären B und $C + (A - A_p)$ in A abgeschlossen, und zwei Punkte von B, C gehörten zu verschiedenen Quasikomponenten. Insbesondere fallen bei end-

[1] Eine zusammenhängende Menge hat eine einzige Quasikomponente, sich selbst.

licher Quasikomponentenzahl (also auch bei endlicher Komponentenzahl) die Quasikomponenten mit den Komponenten zusammen. Gehören $p_1, p_2, \ldots, p_n$ zu paarweise verschiedenen Quasikomponenten, so ist eine Zerlegung $A = A_1 + A_2 + \ldots + A_n$ in lauter relativ abgeschlossene Mengen möglich, wo p_1 zu $A_1, \ldots, p_n$ zu A_n gehört. Denn dies ist für $n = 2$ richtig und überträgt sich von n auf $n+1$: ist p_0 ein $(n+1)^{\text{ter}}$ Punkt, der etwa zu A_1 gehört, und $A = B_0 + B_1$ eine Zerlegung, bei der p_0 zu B_0 und p_1 zu B_1 gehört, so ist $A = \mathfrak{D}(A_1, B_0) + \mathfrak{D}(A_1, B_1) + A_2 + \ldots + A_n$ eine entsprechende Zerlegung für $n+1$.

Ein Beispiel, wo Quasikomponenten und Komponenten nicht identisch sind, ist dieses. R_n sei in der Ebene der durch die Geraden $x = \pm \frac{n}{n+1}$, $y = \pm n$ bestimmte Rechtecksumfang ($n = 1, 2, 3, \ldots$); A die Summe dieser Rechtecke nebst den beiden Geraden $x = \pm 1$. Die Komponenten von A sind die Rechtecke und die beiden einzelnen Geraden. Bei einer Zerlegung $A = A_1 + A_2$ in abgeschlossene Mengen muß einer der Summanden unendlich viele Rechtecke und deren Häufungspunkte, also beide Geraden $x = \pm 1$ enthalten, diese gehören also einer Quasikomponente an.

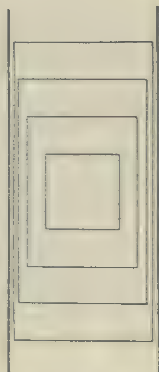


Fig. 6.

§ 8. Dichtigkeit.

Erinnern wir zunächst an ein spezielles Beispiel. Die Menge der rationalen Punkte der Ebene E dringt derart in alle Teile von E ein, daß in jedem noch so kleinen Kreise rationale Punkte liegen. Ein ähnliches Verhalten zeigt sie auch zur Menge der irrationalen Punkte, mit der sie zwar keinen Punkt gemein hat, zwischen deren Punkte sie sich aber überall derart eindrängt, daß in beliebiger Nähe eines irrationalen Punktes stets rationale Punkte liegen. Dies Verhalten, das mit dem Worte Dichtigkeit bezeichnet wird, gibt Anlaß zu folgender allgemeiner Definition:

Die Menge A heißt zu B dicht, wenn in jeder Umgebung eines Punktes von B mindestens ein Punkt von A liegt.

Dies besagt, daß jeder Punkt von B ein α-Punkt von A sein soll; die Bedingung dafür, daß A zu B dicht sei, ist also

(1) $$A_\alpha \geqq B$$

oder die damit offenbar gleichbedeutende

(2) $$A_\alpha \geqq B_\alpha.$$

Wenn A dabei Teilmenge von B ist, so sagen wir: A ist in B dicht. Da dann auch $A_\alpha \leqq B_\alpha$, so wird die Dichtigkeit von A in B durch die Formeln

$$A \leqq B, \quad A_\alpha = B_\alpha$$

dargestellt, von denen die zweite

$$A_\alpha = B_\alpha$$

für sich allein besagt, daß A und B zueinander dicht sind.

Ist A im Raume E dicht ($A_\alpha = E$), so sagt man wohl schlechthin: A ist (absolut) dicht, und unterscheidet davon den allgemeinen Fall als relative Dichtigkeit.

Man erkennt unmittelbar die Richtigkeit der folgenden Sätze:

Jede Menge ist in sich selbst[1] und zu jeder Teilmenge, insbesondere zur Nullmenge dicht. Ist A zu B dicht, so ist auch A zu B_α dicht; A ist in A_α dicht. Ist A zu B dicht, so ist jede Menge $\geqq A$ zu jeder Menge $\leqq B$ dicht. Ist A zu B, B zu C dicht, so ist auch A zu C dicht (transitives Gesetz).

Ferner gilt für beliebig viele Mengen das Summengesetz:

Ist A_1 zu B_1, A_2 zu B_2 usw. dicht, so ist auch $A = \mathfrak{S}(A_1, A_2, \ldots)$ zu $B = \mathfrak{S}(B_1, B_2, \ldots)$ dicht.

Denn es ist

$$A_\alpha \geqq \mathfrak{S}(A_{1\alpha}, A_{2\alpha}, \ldots) \geqq \mathfrak{S}(B_1, B_2, \ldots) = B.$$

Sind A und B zueinander dicht, so wollen wir sie kongruent[2] nennen ($A \equiv B$); diese Relation ist transitiv. Zwei kongruente Mengen haben dieselben Berührungs-, Häufungs- und isolierten Punkte; denn aus $A_\alpha = B_\alpha$ folgt $A_{\alpha\beta} = B_{\alpha\beta}$, $A_\beta = B_\beta$ und $A_\alpha - A_\beta = B_\alpha - B_\beta$, $A_j = B_j$. Ist die eine Menge insichdicht, so auch die andre. Eine insichdichte Menge A ist also in einer perfekten Menge (nämlich A_α) dicht, und umgekehrt ist jede in einer perfekten Menge dichte Menge insichdicht. In jeder Klasse kongruenter Mengen gibt es eine und nur eine abgeschlossene Menge (nämlich in der Klasse, zu der A gehört, die Menge A_α); dies ist die größte Menge der Klasse und alle andern Mengen sind in ihr dicht. Ist A_1 mit B_1, A_2 mit B_2 usw. kongruent, so ist auch $A = \mathfrak{S}(A_1, A_2, \ldots)$ mit $B = \mathfrak{S}(B_1, B_2, \ldots)$ kongruent.

Im Folgenden ist es zweckmäßig, die sonst nicht benutzte Menge

$$A_e = E - A_\alpha = (E - A)_i$$

der äußeren Punkte von A (der inneren des Komplements) heran-

---

[1] „In sich selbst dicht" ist also von „insichdicht" zu unterscheiden (S. 221).

[2] Das Zeichen $\equiv$ für eine symmetrische, transitive Relation ist so bequem, daß man seine mehrfache Anwendung in verschiedenen Bedeutungen nicht beanstanden wird.

zuziehen; ein solcher Punkt hat eine Umgebung, in der kein Punkt von A liegt. A ist zu M dicht, wenn M keinen äußeren Punkt von A enthält.

Wir nennen A zu M undicht, wenn entweder[1] $M = 0$ oder (für $M > 0$) wenn A zu M nicht dicht ist; im letzteren Fall also, wenn M mindestens einen äußeren Punkt von A enthält. Von der Undichtheit sind zwei Verschärfungen bemerkenswert:

A heißt zu M nirgendsdicht, wenn A zu jedem Relativgebiet von M undicht (also zu keinem von Null verschiedenen Relativgebiet von M dicht) ist.

A heißt zu M total undicht, wenn A zu jeder Teilmenge von M undicht (also zu keiner von Null verschiedenen Teilmenge von M dicht) ist.

A ist also zu M nirgendsdicht, wenn in jedem von Null verschiedenen Relativgebiet von M ein äußerer Punkt von A liegt, d. h. wenn $\mathfrak{D}(M, A_e)$ in M dicht ist; A ist zu M total undicht, wenn M nur aus äußeren Punkten von A besteht oder außerhalb A liegt ($M \leqq A_e$).

Zur Nullmenge ist A gleichzeitig dicht, undicht, nirgendsdicht, total undicht.

Hiernach ist das Summengesetz für alle vier Relationen unmittelbar klar:

Ist A zu den sämtlichen Mengen $M_1, M_2, \ldots$ dicht, undicht, nirgendsdicht, total undicht, so ist A auch zu ihrer Summe $M = \mathfrak{S}(M_1, M_2, \ldots)$ dicht, undicht, nirgendsdicht, total undicht.

Das bedarf höchstens für die Relation nirgendsdicht einer Erläuterung: wenn $\mathfrak{D}(M_1, A_e)$ in M_1, $\mathfrak{D}(M_2, A_e)$ in M_2 usw. dicht sind, so ist ihre Summe $\mathfrak{D}(M, A_e)$ in M dicht.

Nach diesem Summengesetz existieren die größten Teilmengen von M, von denen im Folgenden die Rede ist.

Das Verhalten von A zu M hängt von den vier Mengen P, Q, Q', P' ab, die wir der Reihe nach durch die Gleichungen definieren:

(3) $$P = \mathfrak{D}(M, A_\alpha)$$

(4) $$Q = M - P = \mathfrak{D}(M, A_e)$$

(5) $$Q' = \mathfrak{D}(M, Q_\alpha)$$

(6) $$P' = M - Q' = \mathfrak{D}(M, Q_e)$$

P und Q' sind in M abgeschlossen, Q und P' Relativgebiete. Diese Mengen haben folgende Bedeutung:

---

[1] Diese sprachlich etwas anfechtbare Verabredung, eine Menge zur Nullmenge gleichzeitig dicht und undicht zu nennen, dient der formalen Einfachheit.

P ist die größte Teilmenge von M, zu der A dicht ist.

Q ist die größte Teilmenge von M, zu der A total undicht ist.

Q' ist die größte Teilmenge von M, zu der A nirgendsdicht ist.

P' ist das größte Relativgebiet von M, zu dem A dicht ist.

Die ersten beiden Sätze sind evident, da die betreffenden Teilmengen zu A_α resp. A_e gehören müssen.

Der dritte folgt so: soll A zu $M_1 \leqq M$ nirgendsdicht sein, so muß $Q_1 = \mathfrak{D}(M_1, A_e)$ in M_1 dicht sein, also

$$M_1 = \mathfrak{D}(M_1, Q_{1\alpha}) \leqq \mathfrak{D}(M, Q_\alpha) = Q'.$$

Andererseits wird für $M_1 = Q'$

$$Q_1 = \mathfrak{D}(Q', A_e) = \mathfrak{D}(M, Q_\alpha, A_e) = \mathfrak{D}(Q, Q_\alpha) = Q$$

und Q ist in Q' dicht, also A zu Q' nirgendsdicht.

Der vierte Satz wird folgendermaßen bewiesen. Soll $\mathfrak{P}$ ein Relativgebiet von M, also $\mathfrak{Q} = M - \mathfrak{P}$ in M abgeschlossen, und überdies A zu $\mathfrak{P}$ dicht sein, so folgt der Reihe nach

$$\mathfrak{P} \leqq P, \quad \mathfrak{Q} \geqq Q, \quad \mathfrak{Q}_\alpha \geqq Q_\alpha,$$
$$\mathfrak{Q} = \mathfrak{D}(M, \mathfrak{Q}_\alpha) \geqq Q', \ \mathfrak{P} \leqq P';$$

andererseits ist P' selbst ein Relativgebiet und $\leqq P$ (wegen $Q' \geqq Q$), also A zu P' dicht.

Ist A zu M dicht, so ist

$$P = M, \quad Q = 0, \quad P' = M, \quad Q' = 0 \tag{7}$$

und jede dieser vier Gleichungen zieht die übrigen nach sich.

Ist A zu M nirgendsdicht, so ist

$$P' = 0, \quad Q' = M \tag{8}$$

und Q in M dicht.

Ist A zu M total undicht, so ist

$$P = 0, \quad Q = M, \tag{9}$$

was natürlich die Gleichungen (8) zur Folge hat.

Die vorstehenden Betrachtungen gewinnen vielleicht etwas an Anschaulichkeit, wenn man von einer zu M dichten resp. total undichten Menge sagt, daß sie M erfüllt resp. leer läßt. Dann ist P die größte von A erfüllte, Q die größte von A leer gelassene Teilmenge von M, P' das größte von A erfüllte Relativgebiet von M, und wenn der von A leer gelassene Bestandteil Q seinerseits die Menge M erfüllt, so ist A zu M nirgendsdicht.

Ist A in M abgeschlossen, so ist $P = A$, $Q = M - A$; also A in M dicht nur dann, wenn es mit M zusammenfällt, und nirgends-

dicht, wenn sein Komplement $M - A$ in M dicht ist. Ist überdies M der Raum E, so wird einfach $(E = A + B)$

$$P = A, \quad Q = B, \quad Q' = B_\alpha, \quad P' = A_i.$$

Eine abgeschlossene Menge ist also (in E) nirgendsdicht dann und nur dann, wenn sie keine inneren Punkte hat; abgeschlossene nirgendsdichte Mengen sind identisch mit abgeschlossenen Randmengen, also mit Gebietsgrenzen. Denn eine abgeschlossene Randmenge A ist die Grenze des Gebietes $E - A$, und die Grenze eines Gebietes G ist abgeschlossen und der Rand von $E - G$. Ferner sind abgeschlossene nirgendsdichte Mengen die Komplemente dichter Gebiete.

Beispiele. Im euklidischen Raum sind Gebiete insichdichte Mengen, also das Innere jeder Menge in ihrem Kern enthalten $(A_i \leqq A_k)$, und eine separierte abgeschlossene Menge $(A_k = 0)$ ist also gewiß nirgendsdicht. Aber auch abgeschlossene Mengen mit nichtverschwindendem Kern, z. B. perfekte Mengen, können nirgendsdicht sein. Um auf der geraden Linie eine solche zu bilden, verstehen wir unter Δ eine offene Strecke $(a < x < b)$ und unter

$$G = \Delta_1 + \Delta_2 + \dots = \sum \Delta_n$$

eine Summe von abzählbar vielen, paarweise fremden Strecken, also ein lineares Gebiet, das, weil unzusammenhängend, sicher nicht mit der ganzen Geraden identisch ist (§ 9). Damit das Komplement F perfekt werde, vermeiden wir bei der sukzessiven Wahl der Strecken $\Delta_1, \Delta_2, \dots$, daß zwei von

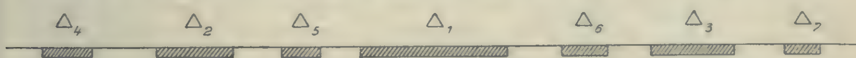


Fig. 7.

ihnen mit einem Endpunkt zusammenstoßen, der ja sonst ein isolierter Punkt von F sein würde; die Figur zeigt den Anfang einer solchen Streckenwahl. Um endlich G dicht, also F nirgendsdicht zu machen, lassen wir G etwa die Menge $R = \{r_1, r_2, \dots\}$ der rationalen Zahlen einschließen; wir wählen alle Δ_n mit irrationalen Endpunkten, bedecken r_1 mit Δ_1, das erste (mit niedrigstem Index behaftete) von Δ_1 nicht eingeschlossene r_β mit Δ_2, das erste von $\Delta_1 + \Delta_2$ nicht eingeschlossene r_γ mit Δ_3 usw. Hiermit wird $R \leqq G$, also G auf der geraden Linie dicht. Wir werden sehen, daß die perfekte Menge F die Mächtigkeit des Kontinuums hat; Mengen so hoher Mächtigkeit können also nirgendsdicht sein. Man kann sogar, wenn δ_n die Länge der Strecke Δ_n bedeutet, die Summe $\sum \delta_n$ (das „Längenmaß" von G) konvergent und beliebig klein machen, was ja mit der angegebenen Konstruktion verträglich ist; Gebiete von beliebig kleinem Längenmaß können die ganze Gerade dicht erfüllen.

Natürlich kann man diese Konstruktion mannigfach abändern. Ein

Beispiel (wo allerdings $\Sigma \delta_n$ divergiert) ist das folgende: die Menge F bestehe aus den Zahlen, die eine Dezimalbruchentwicklung

$$x = x_0, x_1 x_2 \ldots = x_0 + \frac{x_1}{10} + \frac{x_2}{10^2} + \ldots$$

zulassen, in der eine der Ziffern 1—8 nicht vorkommt, etwa die Ziffer 3; d. h. x_0 ist eine beliebige ganze Zahl, aber x_n (für natürliches n) eine der Ziffern 0—9 mit Ausschluß der 3. Das Komplement G besteht aus den Strecken, die von den Endpunkten

$$\begin{array}{lcl} x_0, 3 & \text{und} & x_0, 4 \\ x_0, x_1 3 & \text{und} & x_0, x_1 4 \\ x_0, x_1 x_2 3 & \text{und} & x_0, x_1 x_2 4 \quad \text{usw.} \end{array}$$

begrenzt sind (wieder mit $x_n \neq 3$), diese Endpunkte selbst nicht mitgezählt (z. B. gehört der Punkt $0{,}3 = 0{,}2999\ldots$ zu F); G ist also ein Gebiet und F abgeschlossen, insbesondere perfekt und nirgendsdicht, da jeder Zahl von F sowohl solche Zahlen, die die Ziffer 3 nicht enthalten, als auch solche, die sie enthalten, beliebig nahe kommen, jeder Punkt von F also Häufungspunkt sowohl von F als auch von G ist. F hat die Mächtigkeit $\aleph_0 \cdot 9^{\aleph_0} = 2^{\aleph_0}$ des Kontinuums. Dieselben Betrachtungen kann man natürlich für andere Zifferdarstellungen reeller Zahlen durchführen; das erste Beispiel dieser Art wurde von G. Cantor gegeben, nämlich die Menge der Zahlen, in deren triadischer Entwicklung

$$x = x_0 + \frac{x_1}{3} + \frac{x_2}{3^2} + \cdots$$

die Ziffer 1 nicht vorkommt.

Daß abgeschlossene nirgendsdichte Mengen der Ebene die Mächtigkeit des Kontinuums haben können, ist trivial, da schon die Gebietsgrenzen der Elementargeometrie (z. B. eine gerade Linie oder Kreisperipherie) solche liefern. Man kann aber auch die vorstehenden Betrachtungen auf die Ebene übertragen, indem man unter den Δ_n das Innere von Quadraten, Rechtecken oder dgl. versteht, und so auch ein die Ebene dicht erfüllendes Gebiet von beliebig kleinem Flächenmaß konstruieren.

Wir haben der allgemeinen Theorie noch folgende Sätze hinzuzufügen:

Ist A zu B dicht und G ein beliebiges Gebiet, so ist auch $\mathfrak{D}(A, G)$ zu $\mathfrak{D}(B, G)$ dicht.

Nach den Formeln § 6, (4) ist nämlich

$$\mathfrak{D}(A, G)_\alpha \geqq \mathfrak{D}(A_\alpha, G) \geqq \mathfrak{D}(B, G).$$

Sind A_1, A_2 Relativgebiete von M und in M dicht, so ist auch ihr Durchschnitt $\mathfrak{D}(A_1, A_2)$ in M dicht.

Sei $A_1 = \mathfrak{D}(M, G_1)$, $A_2 = \mathfrak{D}(M, G_2)$, und beide in M dicht. Aus dem vorigen Satze folgt, daß $\mathfrak{D}(A_1, G_2)$ in $\mathfrak{D}(M, G_2)$ oder $\mathfrak{D}(A_1, A_2)$

in A_2 dicht ist; da auch A_2 in M dicht ist, so ist nach dem transitiven Gesetz $\mathfrak{D}(A_1, A_2)$ in M dicht.

Im allgemeinen ist der Durchschnitt zweier in M dichter Mengen keineswegs in M dicht, wie die Mengen der rationalen und der irrationalen Punkte der Ebene lehren.

Sind A_1 und A_2 zu M nirgendsdicht, so ist auch ihre Summe zu M nirgendsdicht.

Bildet man nämlich von A_1, A_2 und $A = \mathfrak{S}(A_1, A_2)$ nach den Formeln (3), (4) die zugehörigen Mengen P_1, P_2, P und Q_1, Q_2, Q, so ist wegen $A_\alpha = \mathfrak{S}(A_{1\alpha}, A_{2\alpha})$

$$P = \mathfrak{S}(P_1, P_2), \quad Q = \mathfrak{D}(Q_1, Q_2).$$

Nun sind nach Annahme Q_1 und Q_2 in M dicht, beide sind ferner Relativgebiete; nach dem vorigen Satz ist also auch Q in M dicht, also A zu M nirgendsdicht.

Die beiden letzten Sätze übertragen sich unmittelbar von zwei auf eine endliche Anzahl von Mengen. Für abzählbar viele Mengen werden sich erst später beschränkte Analoga dieser Sätze ergeben (Kap. VIII. § 9).

Für jede Menge A ist die Menge A_j der isolierten Punkte in der Menge A_s der separierten Punkte dicht. Erinnern wir uns aus § 3 der Zerlegungen

$$A = A_h + A_j = A_k + A_s.$$

Um die Behauptung $A_s \leqq A_{j\alpha}$ zu beweisen, nehmen wir an, x sei kein α-Punkt von A_j; es gebe also eine Umgebung U_x, die keinen isolierten Punkt von A enthält. Die Menge $B = \mathfrak{D}(A, U_x)$ ist dann insichdicht, denn jeder ihrer Punkte ist Häufungspunkt von A, also auch von B; oder wegen $B \leqq A_\beta$ und § 6, (4) ist

$$B_\beta \geqq \mathfrak{D}(A_\beta, U_x) \geqq \mathfrak{D}(B, U_x) = B.$$

Demnach gehört B zu A_k, x ist sicher kein Punkt von A_s; umgekehrt muß also ein Punkt von A_s ein α-Punkt von A_j sein.

Insbesondere ist in einer separierten Menge ($A = A_s$) eine isolierte Menge dicht. Umgekehrt braucht aber eine Menge A, in der eine isolierte Menge dicht ist (welche Menge übrigens notwendig A_j ist, da sie dieselben isolierten Punkte wie A hat), nicht separiert zu sein. Wir können z. B. in der euklidischen Ebene eine isolierte Menge J angeben (etwa die Menge der Punkte mit den rechtwinkligen Koordinaten $x = \frac{m}{n}$, $y = \frac{1}{n}$, wo m eine ganze Zahl und $n = 1, 2, 4, 8, \ldots$ ist), deren Ableitung J_β die Abszissenachse ist; dann ist J in der Menge $A = J + J_\beta = J_\alpha$ dicht, während $A_k = J_\beta$ ist.



Fig. 8.

§ 9. Mengen reeller Zahlen.[1]

Teils als Anwendung des Bisherigen, teils zur Vorbereitung späterer Entwicklungen schalten wir am Ende dieses Kapitels einige Betrachtungen über Mengen reeller Zahlen (oder Punktmengen auf der geraden Linie) ein; auch einige dem Leser bekannte Grundbegriffe der Analysis bedürfen einer mengentheoretischen Beleuchtung.

Die Entfernung zweier reeller Zahlen ist $\overline{xy} = |x - y|$; die damit definierte (sphärische) Umgebung U_x mit dem Radius ϱ ist die Menge der Zahlen y, für die $x - \varrho < y < x + \varrho$. Hiernach sind α-Punkte oder α-Zahlen usw. wie in der allgemeinen Theorie zu definieren. Eine konvergente Menge ist abzählbar, denn ist $x = \lim A$, so ist die Menge A_n der Zahlen von A, für die $\overline{ax} \geqq \frac{1}{n}$, endlich, und A ist die Summe $\mathfrak{S}(A_1, A_2, \ldots)$ mit eventueller Hinzunahme der Zahl x selbst.

Eine Menge A reeller Zahlen heißt nach unten beschränkt, wenn es eine Zahl u gibt, die kleiner ist als alle Zahlen von A $(u < a)$; nach oben beschränkt, wenn es eine Zahl v gibt, die größer ist als alle Zahlen von a $(v > a)$; beschränkt schlechthin, wenn beides der Fall ist, also alle Zahlen von A zwischen zwei festen Zahlen liegen $(u < a < v)$. Die endlichen Mengen sind beschränkt; auch die Nullmenge rechnen wir dazu.

Wir werden zeigen, daß sich hier die Begriffe beschränkt und kompakt decken. Zunächst ist sofort klar, daß eine nicht beschränkte Menge divergente Teilmengen enthält, also nicht kompakt ist; demnach ist eine kompakte Menge beschränkt. Denn ist A etwa nicht nach oben beschränkt, so läßt sich aus ihr eine Zahlenfolge herausgreifen gemäß den Ungleichungen

$$a_2 \geqq a_1 + 1, \quad a_3 \geqq a_2 + 1, \quad \ldots$$

und die Menge $\{a_1, a_2, a_3, \ldots\}$ hat offenbar keine Häufungszahl. Nachher wird sich herausstellen, daß umgekehrt eine beschränkte Menge kompakt ist.

Nach der Dedekindschen Definition der Irrationalzahlen, worüber in Kap. IV, § 5 das Nötige gesagt ist, ist die natürlich geordnete Menge T der reellen Zahlen stetig, d. h. wenn sie in zwei nichtverschwindende Teilmengen $T = U + V$ derart zerlegt wird, daß jede Zahl u (von U) kleiner als jede Zahl v (von V) ist, so hat entweder U eine größte oder V eine kleinste Zahl. Wird diese mit x bezeichnet, so kann x selber ein u oder ein v sein, aber für jedes positive ϱ ist $x - \varrho$ ein u und $x + \varrho$ ein v. Wir sagen, die

---

[1] Vgl. dazu Kap. I, § 11.

Zerlegung $T = U + V$ bestimmt die Zahl x, und nennen eine solche Zerlegung (wo U, V nicht verschwinden) eine eigentliche Zerlegung. Den uneigentlichen Zerlegungen $0 + T$, $T + 0$ ordnet man die uneigentlichen Zahlen $-\infty$, $+\infty$ zu; davon sehen wir aber im folgenden ab.

Wenn in einer Menge A reeller Zahlen eine kleinste Zahl vorhanden ist, so nennen wir sie das Minimum von A und bezeichnen sie mit $\min A$; entsprechend ist $\max A$ zu erklären.

Ist nun A eine Menge reeller Zahlen, so verstehen wir unter $u_\alpha, u_\beta, u_\gamma$ jede reelle Zahl, die keine, resp. nur endlich viele, resp. höchstens abzählbar viele Zahlen a von A übertrifft,[1] und unter $v_\alpha, v_\beta, v_\gamma$ jede reelle Zahl, die das Gegenteil tut, die also mindestens eine, resp. unendlich viele, resp. unabzählbar viele Zahlen a übertrifft. Wenn es solche Zahlen gibt, ist

$$u_\alpha < v_\alpha, \quad u_\beta < v_\beta, \quad u_\gamma < v_\gamma$$

und wir erhalten also drei Zerlegungen

$$T = U_\alpha + V_\alpha = U_\beta + V_\beta = U_\gamma + V_\gamma;$$

indessen braucht es nicht alle zu geben und die Zerlegungen können uneigentliche sein. Soweit es eigentliche Zerlegungen sind, bezeichnen wir die durch sie bestimmten Zahlen mit $\xi_\alpha, \xi_\beta, \xi_\gamma$ und wollen zeigen, daß dies die kleinste α-Zahl, β-Zahl, γ-Zahl oder die kleinste Berührungszahl, Häufungszahl, Verdichtungszahl von A ist, also

$$\xi_\alpha = \min A_\alpha, \quad \xi_\beta = \min A_\beta, \quad \xi_\gamma = \min A_\gamma.$$

Um kurz zu sein, verstehen wir unter $A(x)$ die Menge der Zahlen von A, die kleiner als x sind, unter ϱ eine positive Zahl, unter δ einen der drei Indices α, β, γ und unter k_δ eine der drei Kardinalzahlen

$$k_\alpha = 1, \quad k_\beta = \aleph_0, \quad k_\gamma = \aleph_1.$$

Dann ist $\xi_\delta + \varrho$ ein v_δ, $\xi_\delta - \varrho$ ein u_δ, also $A(\xi_\delta + \varrho)$ von der Mächtigkeit $\geqq k_\delta$, $A(\xi_\delta - \varrho)$ von der Mächtigkeit $< k_\delta$, $A(\xi_\delta + \varrho) - A(\xi_\delta - \varrho)$ von der Mächtigkeit $\geqq k_\delta$. Diese Differenz gehört der Umgebung von ξ_δ mit dem Radius 2ϱ an. Da ϱ beliebig ist, so enthält jede Umgebung von ξ_δ mindestens k_δ Punkte von A, ξ_δ ist ein Punkt von A_δ. Andererseits ist $\xi_\delta - 2\varrho$ kein Punkt von A_δ, denn seine Umgebung mit dem Radius ϱ gehört zu $A(\xi_\delta - \varrho)$ und enthält weniger als k_δ Punkte von A. Da also kein Punkt $< \xi_\delta$ zu A_δ gehört, so ist $\xi_\delta = \min A_\delta$.

---

[1] b übertrifft a, heißt: $b > a$.

Entsprechende Betrachtungen führen zu den drei Zahlen

$$\eta_\alpha = \max A_\alpha, \quad \eta_\beta = \max A_\beta, \quad \eta_\gamma = \max A_\gamma;$$

dabei ist (soweit diese Zahlen existieren) offenbar

$$\xi_\alpha \leqq \xi_\beta \leqq \xi_\gamma \leqq \eta_\gamma \leqq \eta_\beta \leqq \eta_\alpha,$$

weil ja $A_\alpha \geqq A_\beta \geqq A_\gamma$ und stets $\min A \leqq \max A$.

Die Benennungen dieser Zahlen gehen sehr auseinander. Es finden sich für ξ_α die Namen untere Schranke von A (M. Pasch), untere Grenze (K. Weierstrass); für ξ_β untere Grenze (M. Pasch), untere Unbestimmtheitsgrenze (P. du Bois-Reymond), unterer Limes (A. Pringsheim) usw., mit den entsprechenden für η_α, η_β. Wir schließen uns den Bezeichnungen von M. Pasch an: ξ_α die untere Schranke, ξ_β die untere Grenze von A.

Es ist noch die Frage, wann diese Zahlen existieren, wann also, um nur von den ξ zu sprechen, die obigen Zerlegungen eigentliche sind. Die Menge A wird natürlich von Null verschieden angenommen. Zahlen u_α existieren dann und nur dann, wenn A nach unten beschränkt ist, Zahlen v_α existieren stets. Also existiert die untere Schranke ξ_α dann und nur dann, wenn A nach unten beschränkt ist; insbesondere hat eine nach unten beschränkte, abgeschlossene Menge stets ein Minimum $\min A_\alpha = \min A$. Zur Existenz der unteren Grenze ξ_β ist notwendig und hinreichend, daß A nach unten beschränkt und A_β von Null verschieden sei; zur Existenz von ξ_γ ist notwendig und hinreichend, daß A_γ von Null verschieden und nach unten beschränkt sei. Entsprechende Bedingungen (Beschränktheit nach oben) gelten für die η.

Ist A beschränkt und unendlich, so existieren außer ξ_α und η_α auch sicher ξ_β und η_β. Denn dann ist jede Zahl $\xi_\alpha - \varrho$ ($\varrho > 0$) ein u_β und jede Zahl $\eta_\alpha + \varrho$ ein v_β, also existiert ξ_β, und das gleiche gilt von η_β. Das besagt insbesondere:

(Satz von Bolzano-Weierstrass) Jede beschränkte unendliche Menge reeller Zahlen hat mindestens eine Häufungszahl.

Mit andern Worten: jede beschränkte Menge ist kompakt; beschränkte und kompakte Mengen sind hier identisch.

Konvergente Mengen sind demnach beschränkte Mengen mit einer einzigen Häufungszahl; dazu ist die Existenz und Gleichheit von $\xi_\beta = \eta_\beta$ notwendig und hinreichend.

Man schreibt auch

$$\xi_\beta = \liminf A, \quad \eta_\beta = \limsup A$$

und im Falle der Gleichheit $\xi_\beta = \eta_\beta = \lim A$.

Es ist evident, wie alle diese Begriffe auf Zahlenfolgen $\mathfrak{A} = (a_1, a_2, \ldots)$ zu übertragen sind und zu den Zahlen (vgl. § 5)

$$\xi_\alpha = \min \mathfrak{A}_\alpha, \quad \eta_\alpha = \max \mathfrak{A}_\alpha,$$
$$\xi_\beta = \min \mathfrak{A}_\beta, \quad \eta_\beta = \max \mathfrak{A}_\beta$$

führen; man schreibt dann auch

$$\xi_\beta = \lim\inf \mathfrak{A} = \lim\inf a_n, \quad \eta_\beta = \lim\sup \mathfrak{A} = \lim\sup a_n;$$

zur Konvergenz der Folge ist die Existenz und Gleichheit dieser beiden Zahlen

$$\xi_\beta = \eta_\beta = \lim \mathfrak{A} = \lim a_n$$

notwendig und hinreichend.

Den Zusammenhang zwischen diesen Limesbildungen für Zahlenfolgen und denen für Mengenfolgen haben wir bereits in Kap. I, § 11 kennen gelernt, wobei wir dem Leser überlassen wollen, die dort und hier aufgestellten Definitionen für Schranken und Grenzen in Einklang zu bringen.

Wir sind jetzt auch, das erste Mal, in der Lage, zusammenhängende Mengen anzugeben, die nicht bloß aus einem Punkte bestehen, nämlich:

Ein Intervall reeller Zahlen ist stets zusammenhängend.

C sei die Menge der Zahlen x des Intervalls $a \leqq x \leqq b$ $(a < b)$. Sei $C = P + Q$ eine Zerlegung in zwei nichtverschwindende abgeschlossene Mengen; a gehöre zu P, und $q = \min Q$ sei die kleinste Zahl von Q (die Existenz dieses Minimums wurde oben festgestellt). Dann ist $a < q \leqq b$; die Zahlen $a \leqq x < q$ gehören zu P, und q wäre Häufungszahl von P, also P im Widerspruch zur Voraussetzung nicht abgeschlossen. Also ist die angenommene Zerlegung unmöglich, C zusammenhängend.

Nach § 7, I ist eine Menge C reeller Zahlen zusammenhängend, wenn zwei verschiedene Punkte a, b von ihr in einem zu C gehörigen Intervall liegen, wenn also jede Zahl zwischen a und b ebenfalls zu C gehört. Diese Bedingung ist aber auch notwendig, denn wenn x zwischen a und b liegt und nicht zu C gehört, so zerfällt C in zwei Relativgebiete, nämlich die Menge seiner Zahlen $< x$ und $> x$. Danach ist eine zusammenhängende Menge C reeller Zahlen identisch mit einem Stück (Kap. IV, § 4) der natürlich geordneten Zahlenmenge T, und es gilt im Sinne der Addition geordneter Mengen eine Zerlegung

$$T = B + C + D,$$

wobei B und D noch Null sein können oder nicht. Beachtet man

überdies die Stetigkeit von T, so sieht man, daß es nur folgende zusammenhängende lineare Mengen gibt (geometrisch ausgedrückt):

die ganze Gerade T;

Halbgerade mit oder ohne Endpunkt, d. h. die durch

$$x \geqq a, \quad x > a, \quad x \leqq b, \quad x < b$$

definierten Mengen;

Strecken mit beiden, einem oder keinem Endpunkt, d. h. die für $a < b$ durch

$$a \leqq x \leqq b, \quad a < x \leqq b, \quad a \leqq x < b, \quad a < x < b$$

definierten Mengen;

Mengen, die aus einem einzigen Punkt bestehen.

Zusammenhängende lineare Gebiete sind T selbst, die offenen Halbgeraden ($x > a$, $x < b$) und die offenen Strecken ($a < x < b$); zusammenhängende abgeschlossene Mengen sind T, die Halbgeraden mit Endpunkt, die Strecken mit beiden Endpunkten und die einpunktigen Mengen.

Weitere Eigenschaften der linearen Punktmengen werden im folgenden Kapitel bei der Theorie der metrischen, euklidischen Räume usw. zur Sprache kommen.

Achtes Kapitel.

Punktmengen in speziellen Räumen.

§ 1. Gleichwertige Systeme von Umgebungen.

Wenn in der Menge E nebst den Umgebungen U_x noch ein zweites System von Umgebungen V_x vorhanden ist, das den vier Axiomen (A) bis (D) in Kap. VII, § 1 genügt, so bleiben die Entwicklungen des vorigen Kapitels auch für diese neuen Umgebungen in Kraft, aber die neuen Mengen $A_{(\alpha)}$ usw. brauchen mit den alten A_α usw. natürlich nicht identisch zu sein. Sollte aber, und zwar für jede Menge A, $A_{(\alpha)} = A_\alpha$ sein, so lehrt die in Kap. VII, § 3 am Ende besprochene Zurückführung aller Begriffe auf den Grundbegriff des α-Punktes, daß dann auch $A_{(\beta)} = A_\beta$, $A_{(\gamma)} = A_\gamma$, $A_{(k)} = A_k$ usw. ist, daß also die Resultate des vorigen Kapitels nicht nur formal richtig, sondern auch dem Inhalt nach unverändert bleiben. Wir wollen die beiden Umgebungssysteme in diesem Fall gleichwertig nennen. Die Bedingung dafür läßt sich leicht in folgende Form

bringen: die beiden Systeme von Umgebungen U_x, V_x sind gleichwertig dann und nur dann, wenn jedes U_x ein V_x und jedes V_x ein U_x als Teilmenge enthält. In der Tat, wenn es zu jedem V_x ein $U_x \leqq V_x$ gibt, so ist $A_\alpha \leqq A_{(\alpha)}$; denn ist x ein α-Punkt von A, so enthält jedes U_x, also auch jedes V_x mindestens einen Punkt von A, und x ist auch (α)-Punkt von A. Gibt es umgekehrt zu jedem U_x ein $V_x \leqq U_x$, so ist $A_\alpha \geqq A_{(\alpha)}$, und ist beides der Fall, so ist $A_\alpha = A_{(\alpha)}$. Die Bedingung ist also hinreichend; daß sie auch notwendig ist, sieht man folgendermaßen. Gibt es nicht zu jedem V_x ein $U_x \leqq V_x$, so sei V ein solches V_x, zu dem es kein $U_x \leqq V$ gibt. Für jedes U_x ist also $\mathfrak{D}(U_x, V) \neq U_x$, $\mathfrak{D}(U_x, E - V)$ von Null verschieden; das heißt aber, daß x ein α-Punkt von $E - V$ ist, während es kein (α)-Punkt dieser Menge ist. Sobald also unsere Bedingung nicht erfüllt ist, gibt es Mengen, für die $A_{(\alpha)} \neq A_\alpha$.

Die genannte Forderung, mit den Umgebungsaxiomen für die V_x vereinigt, führt unmittelbar zu der folgenden notwendigen und hinreichenden Gleichwertigkeitsbedingung: die V_x müssen (nach der alten Definition auf Grund der U_x) Gebiete sein, die den Punkt x enthalten, und jedes U_x muß ein V_x als Teilmenge enthalten.

Wir geben einige Beispiele.

(1) Im euklidischen Raume E_n sei V_x die Menge der Punkte y, die für ein positives ϱ den Ungleichungen

$$|x_1 - y_1| < \varrho, \ |x_2 - y_2| < \varrho, \ \ldots, \ |x_n - y_n| < \varrho$$

genügen, also das Innere eines n-dimensionalen Würfels mit dem Mittelpunkt x und Kanten parallel den Koordinatenachsen. Diese kubischen Umgebungen sind mit den sphärischen U_x gleichwertig; sie können übrigens wie diese durch Entfernungen definiert werden, indem man als Entfernung $\overline{xy}$ das Maximum der n Größen $|x_i - y_i|$ $(i = 1, 2, \ldots, n)$ festsetzt. Ähnlich verhalten sich ellipsoidische Umgebungen u. dgl.

(2) In E_n sei V_x die Menge der Punkte y, die für ein positives ϱ den Bedingungen

$$|x_1 - y_1| < \varrho, \ x_2 = y_2, \ \ldots, \ x_n = y_n$$

genügen, also das Innere einer zur ersten Koordinatenachse parallelen Strecke. Da (für $n > 1$) diese V_x keine Gebiete im Sinne der gewöhnlichen Definition sind, so sind sie mit den sphärischen Umgebungen U_x nicht gleichwertig. In der Tat hat hier z. B. eine der zweiten Koordinatenachse parallele Gerade A keinen Häufungspunkt, also $A_{(\beta)} = 0$, während $A_\beta = A$ ist.

Von allgemeinerem Charakter und für das Folgende sehr wichtig sind die beiden nächsten Beispiele.

(3) Ist E ein metrischer Raum, U_x eine sphärische Umgebung, so verstehe man unter V_x eine sphärische Umgebung mit rationalem Radius. In diesem System der V_x, das mit dem der U_x gleichwertig ist, ist die Menge der Umgebungen eines bestimmten Punktes x höchstens abzählbar.

(4) Ist E ein metrischer Raum, in dem eine abzählbare Menge R dicht ist (wie im euklidischen die Menge der rationalen Punkte), so betrachten wir jetzt nur die Kugeln U_r mit einem zu R gehörigen Mittelpunkte r und mit rationalem Radius ϱ; als Umgebungen V_x eines Punktes x mögen dann alle diejenigen U_r bezeichnet werden, die den Punkt x enthalten. Wir zeigen, daß diese V_x mit den U_x gleichwertig sind. Zunächst ist jedes V_x ein Gebiet, das x enthält. Um sodann zu beweisen, daß jedes U_x ein V_x enthält, verstehen wir unter U_x eine Kugel mit dem Mittelpunkt x und dem Radius ξ. Es gibt dann einen Punkt r, für den $\overline{xr} < \frac{1}{2}\xi$. Ist ferner ϱ eine rationale Zahl derart, daß $\overline{xr} < \varrho < \frac{1}{2}\xi$, und U_r die Kugel mit dem Mittelpunkt r und dem Radius ϱ, so enthält U_r den Punkt x ($\overline{rx} < \varrho$) und ist in U_x enthalten, denn aus $\overline{ry} < \varrho$ folgt

$$\overline{xy} \leqq \overline{xr} + \overline{ry} < 2\varrho < \xi.$$

Dieses U_r ist also ein $V_x \leqq U_x$.

In diesem Beispiel ist nicht nur, wie im vorigen, bei festem x die Menge aller V_x, sondern die Menge aller verschiedenen V höchstens abzählbar; übrigens ist sie genau abzählbar, da nach Kap. VII, § 2, (6) ein Raum mit endlich vielen Umgebungen auch nur endlich viele Punkte hat (und umgekehrt), während hier $E \geqq R$ unendlich angenommen ist.

Wir werden uns nun dem gewöhnlichen Raume und den ihm nächstverwandten Mengen um einen großen Schritt nähern, wenn wir im Einklang mit den letzten Beispielen zu den Voraussetzungen des vorigen Kapitels die hinzunehmen, daß das System der Umgebungen U_x mit einem System von Umgebungen V_x gleichwertig sei, das entweder als Ganzes abzählbar ist (Beispiel (4)) oder wenigstens jedem einzelnen Punkte höchstens abzählbar viele Umgebungen zuordnet (Beispiel (3)). Da wir uns an eine bestimmte Bedeutung der U_x nicht gebunden haben, so können wir dem System der U_x selbst die soeben genannten Eigenschaften beilegen. Wir ziehen vor, zunächst die allgemeinere, sodann die speziellere Eigenschaft für sich zu postulieren und ihre Konsequenzen zu entwickeln, also die Beziehungen zum Abzählbaren, die wir dem Raume aufzuzwingen im Begriff sind, in zwei Stufen zu trennen. Indem wir im übrigen also die Umgebungstheorie (ohne metrische Spezialisierung) noch festhalten, unterscheiden wir die zwei

Abzählbarkeitsaxiome:

(E) **Für jeden Punkt x ist die Menge seiner verschiedenen Umgebungen U_x höchstens abzählbar.**

(F) **Die Menge aller verschiedenen Umgebungen U ist abzählbar.**

§ 2. Das erste Abzählbarkeitsaxiom.

Die Konsequenzen dieses schwächeren Axioms betreffen hauptsächlich konvergente Mengen und Folgen, die erst jetzt eine wichtigere Rolle zu spielen beginnen.

I. Jede konvergente Menge ist abzählbar.

Die Menge der verschiedenen Umgebungen eines Punktes x ist endlich oder abzählbar; um beide Fälle im Folgenden nicht immer unterscheiden zu müssen, denken wir uns diese Umgebungen in Gestalt einer Folge $V_1, V_2, \ldots$ gebracht, indem wir eventuell eine von ihnen unendlich oft schreiben. Der Durchschnitt aller dieser Umgebungen besteht aus x allein (S. 217), es ist also

$$\{x\} = \mathfrak{D}(V_1, V_2, \ldots),$$
$$E - \{x\} = \mathfrak{S}(E - V_1, E - V_2, \ldots).$$

Schneiden wir dies mit A und ist $\lim A = x$, so kommt linkerhand A oder $A - \{x\}$, rechterhand die Summe der Durchschnitte $\mathfrak{D}(A, E - V_n)$, deren jeder nach der Definition der Konvergenz endlich ist, also eine höchstens abzählbare Menge. A ist also höchstens abzählbar und als unendliche Menge genau abzählbar.

Ferner gilt der wichtige Satz von der Trennbarkeit der Häufungspunkte:

II. Ist x ein Häufungspunkt von B, so hat B eine nach x konvergente Teilmenge A ($\lim A = x$).

Sind wieder $V_1, V_2, \ldots$ die Umgebungen von x, so setzen wir

$$G_n = \mathfrak{D}(V_1, V_2, \ldots, V_n) \qquad (G_1 \geqq G_2 \geqq \ldots).$$

Jedes G_n ist ein Gebiet und enthält x, also auch eine Umgebung von x und demnach unendlich viele Punkte von B; man wähle unter diesen einen Punkt a_n so, daß der Reihe nach $a_1, a_2, a_3, \ldots$ paarweise verschieden sind; A sei die Menge dieser Punkte. Jede Umgebung von x enthält fast alle Punkte von A, denn V_n enthält $a_n, a_{n+1}, a_{n+2}, \ldots$; also ist $x = \lim A$.

Ebenso hat eine Folge mit dem Häufungspunkt x eine nach x konvergente Teilfolge.

Beiläufig folgt daraus, daß es sicher unendliche isolierte Mengen gibt, falls der ganze Raum E unendlich ist. Denn entweder (was nach den bisherigen Voraussetzungen nicht ausgeschlossen wäre) ist E selbst isoliert; oder es sei x ein Punkt von E_β und A eine wie soeben konstruierte, nach x konvergente Menge, die den Punkt x

selbst nicht enthält. Diese ist isoliert, weil sie ihren einzigen Häufungspunkt nicht enthält.

Aus II folgt auch, daß eine in sich kompakte Menge B (d. h. von der jede unendliche Teilmenge mindestens einen Häufungspunkt in B hat) abgeschlossen ist. Denn ist x ein Häufungspunkt von B, A eine nach x konvergente Teilmenge, so muß A einen Häufungspunkt in B haben, und da x der einzige Punkt von A_β ist, so gehört x zu B. Abgeschlossene kompakte Mengen und in sich kompakte Mengen sind identisch.

Wir wollen hier einmal ein Gegenbeispiel geben, wonach der Satz II ohne das erste Abzählbarkeitsaxiom nicht zu gelten braucht. E sei ein euklidischer Raum und U_x seien die gewöhnlichen sphärischen Umgebungen: wir definieren neue Umgebungen V_x, als die Mengen, die aus den U_x durch Weglassung endlich oder abzählbar vieler Punkte bestehen (aber den Punkt x enthalten). Auch die V_x erfüllen die Umgebungsaxiome. Offenbar ist nun ein Häufungspunkt auf Grund der neuen Umgebungen ein Verdichtungspunkt im gewöhnlichen Sinne und umgekehrt; in § 3 werden wir sehen, daß eine Menge des euklidischen Raumes entweder keinen oder unabzählbar viele Verdichtungspunkte hat; Mengen mit einem einzigen Häufungspunkt und insbesondere konvergente Mengen existieren also auf Grund der augenblicklichen Verabredung nicht.

Weiter beweisen wir:

III. Die α-Punkte einer Menge sind identisch mit den Limites konvergenter Folgen von Punkten der Menge.

Ist nämlich B eine beliebige Menge, $\mathfrak{A}$ eine konvergente Folge von Punkten aus B, so ist $x = \lim \mathfrak{A}$ ein α-Punkt von B, wie unmittelbar klar ist. Umgekehrt, ein α-Punkt x von B ist als Limes einer solchen Folge $\mathfrak{A}$ darstellbar. Denn gehört x zu B selbst, so ist er Limes der Folge $(x, x, x, \ldots)$; gehört x nicht zu B, also zu B_β, so gibt es nach II eine nach x konvergente, also nach I abzählbare Teilmenge $A \leqq B$, und schreibt man diese als Folge $\mathfrak{A}$, so ist $x = \lim \mathfrak{A}$.

Wir wollen dem Leser überlassen, sich klar zu machen, daß die β-Punkte von B die Limites konvergenter Folgen von paarweise verschiedenen Punkten der Menge B sind.

Auch die abgeschlossenen Limites einer Mengenfolge (Kap. VII, § 5) können jetzt mit Hilfe konvergenter Folgen erklärt werden: nämlich $\overline{L} = \overline{\mathrm{Lim}} \inf A_n$ als Menge aller Punkte $\lim \mathfrak{P}$ und $\overline{M} = \overline{\mathrm{Lim}} \sup A_n$ als Menge aller Punkte $\lim \mathfrak{Q}$. Dabei bedeute wie damals $P = \{p_1, p_2, \ldots\}$ eine Folge aus fast allen, $Q = \{q_1, q_2, \ldots\}$ eine Folge aus unendlich vielen natürlichen Zahlen,

$$\mathfrak{P} = (a_{p_1}, a_{p_2}, \ldots) \quad \text{und} \quad \mathfrak{Q} = (a_{q_1}, a_{q_2}, \ldots)$$

entsprechende konvergente Punktfolgen, a_n einen Punkt aus A_n. In der Tat sieht man unmittelbar, daß für $x = \lim \mathfrak{P}$ resp. $x = \lim \mathfrak{Q}$ jede Umgebung von x Punkte mit fast allen resp. unendlich vielen Mengen A_n gemein hat, daß x also zu $\overline{L}$ resp. zu $\overline{M}$ gehört. Um auch die Umkehrung zu beweisen, bezeichnen wir wieder die sämtlichen Umgebungen von x mit $V_1, V_2, \dots$ und führen die Gebiete G_n ein wie beim Beweise von II. Gehört nun erstens x zu $\overline{L}$, so hat jedes G_m Punkte mit fast allen A_n gemein, d. h. für $n \geqq n_m$, wobei wir $n_1 < n_2 < \dots$ annehmen dürfen. Demgemäß sei, für $n_m \leqq n < n_{m+1}$, a_n ein Punkt von $\mathfrak{D}(G_m, A_n)$. Von der Folge

$$\mathfrak{P} = (a_{n_1}, a_{n_1+1}, a_{n_1+2}, \dots)$$

enthält dann $V_m \geqq G_m$ alle Punkte a_n für $n \geqq n_m$, also ist $x = \lim \mathfrak{P}$. Gehört zweitens x zu $\overline{M}$, so enthält jedes G_m Punkte von unendlich vielen A_n, und wir können die Zahlen $q_1 < q_2 < q_3 < \dots$ so bestimmen, daß G_m mindestens einen Punkt a_{q_m} von A_{q_m} enthält. Von der Folge

$$\mathfrak{Q} = (a_{q_1}, a_{q_2}, a_{q_3}, \dots)$$

enthält dann V_m alle Punkte a_n für $n \geqq q_m$, also $x = \lim \mathfrak{Q}$.

Wenn alle A_n von 0 verschieden sind, können wir auch sagen: $\overline{L}$ ist die Menge aller Limespunkte und $\overline{M}$ die Menge aller Häufungspunkte von Folgen $(a_1, a_2, a_3, \dots)$.

Wir haben in diesem Paragraphen die Wichtigkeit der konvergenten Folgen erkannt. Hier vereinigt sich auch unsere Darstellung mit der auf den Limesbegriff (vgl. Kap. VII, § 1) gegründeten allgemeinen Mengentheorie von M. Fréchet, der wir einige Worte zu widmen haben. In dieser Theorie ist anfänglich weder von Entfernungen noch von Umgebungen die Rede, sondern die Folgen $\mathfrak{A} = (a_1, a_2, \dots)$ aus Elementen (Punkten) der Menge E sind die Quelle aller Beziehungen. Es bestehe dann eine Funktion $x = f(\mathfrak{A})$, die gewissen (nicht allen) Folgen $\mathfrak{A}$ Punkte x von E zuordnet. Eine solche Folge, der ein Punkt zugeordnet ist, heiße konvergent, der zugeordnete Punkt $x = f(\mathfrak{A})$ heiße der Limes von $\mathfrak{A}$ und werde mit

$$x = \lim \mathfrak{A} = \lim a_n$$

bezeichnet; und diese Zuordnung erfülle die beiden Konvergenzaxiome:

($\mathfrak{A}$) Eine aus gleichen Punkten bestehende Folge $(a, a, a, \dots)$ ist konvergent und hat den Limes a.

($\mathfrak{B}$) Jede Teilfolge $(a_{n_1}, a_{n_2}, \dots)$ einer konvergenten Folge $(a_1, a_2, \dots)$, wo also $n_1, n_2, \dots$ wachsende natürliche Zahlen sind, ist wieder konvergent und hat denselben Limes.

Wir definieren, indem wir uns übrigens unserer eigenen Formelsprache bedienen, zu einer Menge $A \leqq E$ wieder die α-, β-, γ-Punkte:

x heißt ein α-Punkt von A, wenn er Limes einer konvergenten, aus Punkten von A bestehenden Folge ist.

x heißt ein β-Punkt von A, wenn er α-Punkt jeder Teilmenge A' von A ist, die alle Punkte von A bis auf endlich viele enthält ($A - A'$ endlich).

x heißt ein γ-Punkt von A, wenn er α-Punkt jeder Teilmenge A'' von A ist, die alle Punkte von A bis auf höchstens abzählbar viele enthält ($A - A''$ höchstens abzählbar).[1]

Unter den Folgen, die nach einem β-Punkt (Häufungspunkt) x von A konvergieren, müssen sich welche befinden, die unendlich viele verschiedene Punkte enthalten. Denn gehört x nicht zu A, so ist $a_n \neq x$; es können dann nicht unendliche viele a_n gleich sein ($a_{n_1} = a_{n_2} = \ldots = a$), weil sonst nach ($\mathfrak{A}$), ($\mathfrak{B}$) $\lim a_n = a \neq x$ wäre, und die Folge enthält also unendlich viele verschiedene a_n. Gehört x aber zu A, so soll x ja auch α-Punkt von $A - \{x\}$ sein, und der Schluß vollendet sich wie soeben. Unter Beachtung von ($\mathfrak{B}$) ist also ein Häufungspunkt als $x = \lim a_n$ mit paarweise verschiedenen a_n darstellbar, und umgekehrt ist ein solcher Punkt stets Häufungspunkt. Ein Häufungspunkt kann demgemäß als Limes einer konvergenten, aus paarweise verschiedenen Punkten von A bestehenden Folge erklärt werden (so Fréchet, der den Begriff α-Punkt nicht benutzt). Die vorstehende Überlegung enthält zugleich den Beweis der Gleichung

$$A_\alpha = \mathfrak{S}(A, A_\beta).$$

Ein γ-Punkt kann auch als β-Punkt jeder Menge A'' ($A - A''$ höchstens abzählbar) erklärt werden.

Eine Menge heißt abgeschlossen, wenn $A \geqq A_\beta$ oder $A = A_\alpha$, usw.

Wir wollen die weitere Entwicklung dieser Theorie hier nicht verfolgen. In gewisser Beziehung ist unsere Umgebungstheorie allgemeiner als die Limestheorie, da wir zunächst keine Beziehung zum Abzählbaren voraussetzen (die erst durch das Axiom (E) hineinkommt); in anderer Hinsicht steht es wieder umgekehrt, da die Limesaxiome ($\mathfrak{A}$), ($\mathfrak{B}$) noch sehr wenig aussagen und einen großen Spielraum für weitere Möglichkeiten offen lassen. In dieser Beziehung möge ein erheblicher Unterschied bemerkt werden: auf Grund der beiden Axiome ($\mathfrak{A}$), ($\mathfrak{B}$) allein braucht eine Menge A_α nicht abgeschlossen zu sein. Dafür gibt es ein interessantes Beispiel, auf das wir später noch zurückkommen werden.

---

[1] In Kap. VII, § 3 haben wir zum Schluß die β- und γ-Punkte in gleicher Weise auf die α-Punkte zurückgeführt.

Als Elemente (Punkte) figurieren die reellen Funktionen $x = x(t)$ einer reellen Variablen t. Die Definition des Limes sei

$$x = \lim x_n,$$

wenn für jedes einzelne t im gewöhnlichen Sinne der Konvergenz von Zahlenfolgen

$$x(t) = \lim x_n(t).$$

Ist nun etwa F die Menge der stetigen Funktionen $f(t)$, so ist $G = F_\alpha$ die Menge der durch konvergente Folgen stetiger Funktionen definierten Limites

$$g(t) = \lim f_n(t)$$

und $H = G_\alpha = F_{\alpha\alpha}$ die Menge der durch konvergente Folgen von Funktionen $g(t)$ definierten

$$h(t) = \lim g_n(t).$$

Die Funktionen $g(t)$ können nun zwar unstetig sein, indessen, wie wir später (Kap. IX, § 4) sehen werden, haben sie immer auch Stetigkeitspunkte. Dagegen können die Funktionen $h(t)$ bereits überall unstetig sein, und daher ist $H > G$, also $G = F_\alpha$ nicht abgeschlossen. Z. B. ist die Funktion

$$g(t) = \lim \frac{1}{1 + n t^2}$$

überall Null bis auf $g(0) = 1$; bringt man dann die rationalen Zahlen in eine Folge $r_1, r_2, \ldots$, so ist die Funktion

$$g_n(t) = g(t - r_1) + g(t - r_2) + \ldots + g(t - r_n)$$

immer noch dem System G angehörig, dagegen die Funktion

$$h(t) = \lim g_n(t) = g(t - r_1) + g(t - r_2) + \ldots$$

an allen rationalen Stellen Eins, an den irrationalen Null, also überall unstetig. Eine andere Darstellung (Dirichlet) derselben Funktion als Limes von Limites stetiger Funktionen ist

$$h(t) = \lim_m \lim_n (\cos m!\,\pi t)^{2n}.$$

Es ist lehrreich, als Gegenstück denselben Fall mit anderer Definition des Limes zu betrachten. Definieren wir, indem wir der Einfachheit halber nur beschränkte Funktionen $x = x(t)$ zulassen, als Entfernung $\overline{xy}$ die obere Schranke[1] von $|x(t) - y(t)|$, so sind die Entfernungsaxiome (S. 211) erfüllt und die damit definierten Umgebungen U_x erfüllen die Umgebungsaxiome. Wenn wir uns dabei auf rationale „Radien" ϱ (in der Bedingung $\overline{xy} < \varrho$) beschränken, so gilt auch das erste Abzählbarkeitsaxiom (E) und der Satz III, daß

[1] Vgl. dazu Kap. VII, § 9. Die obere Schranke einer Funktion ist die obere Schranke der Menge der Funktionswerte.

die Elemente von A_α die Limites konvergenter Folgen von Elementen aus A sind. Die Konvergenzbedingung ist jetzt aber eine andere als soeben: es ist

$$x = \lim x_n,$$

wenn in jeder Umgebung $\overline{xy} < \varrho$ fast alle x_n liegen, und dazu ist notwendig und hinreichend, daß für jede positive Zahl ϱ von einem gewissen Index an

$$|x_n(t) - x(t)| < \varrho$$

ist für alle t. Das ist nun mehr als bloße Konvergenz von $x_n(t)$ nach $x(t)$ für jedes einzelne t, nämlich gleichmäßige Konvergenz. Ist jetzt F die Menge der stetigen Funktionen $f(t)$, so ist F_α die Menge der durch gleichmäßig konvergente Folgen stetiger Funktionen definierten Limesfunktionen, und diese sind bekanntlich wieder stetig (vgl. auch Kap. IX, § 4); es ist also $F_\alpha = F$, demgemäß F_α abgeschlossen, wie es ja in der Umgebungstheorie sein muß.

§ 3. Das zweite Abzählbarkeitsaxiom.

Wir prüfen jetzt die Tragweite des stärkeren Axioms (F), wonach die Menge aller verschiedenen Umgebungen U nur abzählbar ist. Wir bezeichnen sie demgemäß mit $U_1, U_2, \ldots, U_n, \ldots$. Nach wie vor bedeute U_x eine Umgebung des Punktes x (es brauchen nicht alle U_n, die x enthalten, als Umgebungen von x definiert zu sein). Die Menge der verschiedenen U_x ist auf Grund von (F) höchstens abzählbar, (E) also eine Folge von (F). Wir bemerken, daß mit Zulassung von (F) der ganze Raum E als unendliche Menge angenommen ist.

Zunächst beginnen die Verdichtungspunkte oder γ-Punkte, die wir bisher nur formal mit berücksichtigt haben, ohne etwas Positives über sie zu wissen, eine wesentliche Rolle zu spielen. Wir sehen jetzt: wenn kein Punkt der Menge A ein γ-Punkt ist ($\mathfrak{D}(A, A_\gamma) = 0$), so läßt sich zu jedem Punkt a der Menge A eine Umgebung U_a angeben, die höchstens abzählbar viele Punkte von A enthält. Die Menge aller verschiedenen U_a ist höchstens abzählbar und ihre Summe enthält also auch nur höchstens abzählbar viele Punkte von A. Da aber A ganz in dieser Summe enthalten ist, so ist A selber höchstens abzählbar. Während wir also bisher nur wußten, daß für ein höchstens abzählbares A die Menge A_γ verschwindet, erhalten wir jetzt die (verschärfte) Umkehrung:

I. Wenn $\mathfrak{D}(A, A_\gamma) = 0$, so ist A höchstens abzählbar.

Eine unabzählbare Menge hat also sicher, sogar unter ihren eigenen Punkten, Verdichtungspunkte; als Gegenstück dazu haben

wir bewiesen (§ 2, II), daß es sicher unendliche Mengen mit $\mathfrak{D}(A, A_\beta) = 0$ gibt. Ohne das Axiom (F) ist der Satz I nicht richtig, wie das Beispiel (2) in § 1 zeigt, wo eine ganze Gerade nicht einmal einen Häufungspunkt hat; die dort definierten Umgebungen sind mit keinem abzählbaren System gleichwertig.

Wir setzen nunmehr

$$A_v = \mathfrak{D}(A, A_\gamma), \quad A = A_v + A_u,$$

sodaß A_v die Menge der zu A gehörigen Verdichtungspunkte von A, A_u die Menge der übrigen, unverdichteten Punkte von A ist. Wir haben jetzt das Verhältnis der drei Zerlegungen

(1) $$A = A_h + A_j, \quad A_h = \mathfrak{D}(A, A_\beta),$$

(2) $$A = A_k + A_s,$$

(3) $$A = A_v + A_u, \quad A_v = \mathfrak{D}(A, A_\gamma)$$

zu betrachten, wo A_h die Menge der zu A gehörigen Häufungspunkte, A_j die der isolierten Punkte, A_k der Kern oder die größte insichdichte Teilmenge von A, A_s die Menge der separierten Punkte ist (Kap. VII, § 3).

Da die Punkte von A_u keine Verdichtungspunkte von A, also auch nicht von A_u sind, so ist nach I A_u höchstens abzählbar. Daraus folgt $A_{u\gamma} = 0$, also nach (3)

$$A_\gamma = \mathfrak{S}(A_{v\gamma}, A_{u\gamma}) = A_{v\gamma},$$

$$A_v \leqq A_\gamma = A_{v\gamma} \leqq A_{v\beta},$$

oder A_v ist insichdicht. Da A_k die größte insichdichte Teilmenge von A und nach früheren Betrachtungen $\leqq A_h$ war, so folgt

$$A_h \geqq A_k \geqq A_v,$$

also

$$A_j \leqq A_s \leqq A_u,$$

und demgemäß sind die letzten drei Mengen alle höchstens abzählbar. Das gibt den Satz:

II. Die Menge der isolierten, der separierten und der unverdichteten Punkte einer Menge A ist höchstens abzählbar. Die Menge der zu A gehörigen Verdichtungspunkte von A ist insichdicht, also Teilmenge des Kerns von A.

Insbesondere sind isolierte, separierte und unverdichtete Mengen höchstens abzählbar. Ist A unabzählbar, so sind auch A_v, A_k, A_h, A_γ, A_β, A_α unabzählbar. Mengen mit höchstens abzählbarer Ableitung sind höchstens abzählbar; konvergente und divergente Mengen sind abzählbar (für die konvergenten ist uns dies schon bekannt).

Man bemerkt, wie der im vorigen Kapitel herrschende formale Parallelismus zwischen α-, β- und γ-Punkten durch die Abzählbar-

keitsaxiome aufgehoben wird. Es gibt im Unabzählbaren kein Gegenstück zu divergenten und konvergenten Mengen, d. h. es gibt keine unabzählbare Menge ohne Verdichtungspunkt oder mit einem einzigen Verdichtungspunkt.

Wir können nun auch die Iterationsformeln § 3, (5) des vorigen Kapitels vervollständigen. Wenn eine Menge in der Menge ihrer Verdichtungspunkte enthalten ist ($B \leqq B_\gamma$), so ist, weil B_γ abgeschlossen ist,

$$B_\gamma \leqq B_{\gamma\gamma} \leqq B_{\gamma\beta} \leqq B_{\gamma\alpha} = B_\gamma,$$

also

$$B_\gamma = B_{\gamma\alpha} = B_{\gamma\beta} = B_{\gamma\gamma}.$$

Nun trifft dies wegen $A_v \leqq A_\gamma = A_{v\gamma}$ auf $B = A_v$ zu, und da hierfür $B_\gamma = A_\gamma$, so haben wir für jede beliebige Menge A

$$A_\gamma = A_{\gamma\alpha} = A_{\gamma\beta} = A_{\gamma\gamma},$$

welche Gleichungen besagen, daß A_γ perfekt ist.

Aus der Gleichung

$$A_\alpha = \mathfrak{S}(A, A_\beta)$$

folgt noch, indem man die Menge der β- und γ-Punkte nimmt, die schon früher bekannte Gleichung

$$A_{\alpha\beta} = \mathfrak{S}(A_\beta, A_{\beta\beta}) = A_\beta \qquad (\text{wegen } A_{\beta\beta} \leqq A_{\beta\alpha} = A_\beta)$$

und die neue

$$A_{\alpha\gamma} = \mathfrak{S}(A_\gamma, A_{\beta\gamma}) = A_{\beta\gamma} \qquad (\text{wegen } A_\gamma = A_{\gamma\gamma} \leqq A_{\beta\gamma}).$$

Wir erhalten damit die folgende Zusammenstellung der neun Mengen $A_{\alpha\alpha} \dots A_{\gamma\gamma}$, in der sich nur $A_{\beta\beta}$ und $A_{\alpha\gamma} = A_{\beta\gamma}$ im allgemeinen nicht durch Mengen mit einem Index ausdrücken lassen:

$$(4) \qquad \begin{cases} A_{\alpha\alpha} = A_\alpha, & A_{\alpha\beta} = A_\beta, & A_{\alpha\gamma} = A_{\beta\gamma}, \\ A_{\beta\alpha} = A_\beta, & A_{\beta\beta} = A_{\beta\beta}, & A_{\beta\gamma} = A_{\alpha\gamma}, \\ A_{\gamma\alpha} = A_\gamma, & A_{\gamma\beta} = A_\gamma, & A_{\gamma\gamma} = A_\gamma. \end{cases}$$

Weitere Wiederholung würde, wie man hieraus unmittelbar sieht, außer den höheren Ableitungen $A_{\beta\beta\beta}$ usw. keine neuen Mengen liefern. Jede Menge enthält die darunter und die rechts stehende als Teilmenge, also

$$A_\alpha \geqq A_\beta \geqq A_{\beta\beta} \geqq A_{\alpha\gamma} = A_{\beta\gamma} \geqq A_\gamma.$$

Weitere Schlüsse aus (F) beziehen sich auf Gebiete und abgeschlossene Mengen.

III. Eine Menge paarweise fremder Gebiete ist höchstens abzählbar.

Denn sehen wir von dem einen, der Menge eventuell angehörigen Gebiet Null ab, so können wir jedem G der Menge ein $U_n \leqq G$ zuordnen, diese U_n sind paarweise verschieden und ihre Menge höchstens abzählbar.

Ferner sahen wir in Kap. VII, § 2, (7), daß sich jedes von 0 verschiedene Gebiet als Summe der in ihm enthaltenen Umgebungen darstellen läßt; das gibt hier die Formel

(5) $$G = \mathfrak{S}(U_m, U_n, \ldots),$$

wo die $m, n, \ldots$ natürliche Zahlen sind. Diese Darstellung, worin alle in G enthaltenen Umgebungen aufgenommen sind, ist übrigens nicht die einzige. Hiernach entspricht jeder (von 0 verschiedenen) Menge natürlicher Zahlen $m, n, \ldots$ ein (von 0 verschiedenes) Gebiet, umgekehrt jedem Gebiet eine oder mehrere Mengen natürlicher Zahlen. Die Menge aller Mengen natürlicher Zahlen hat die Mächtigkeit $2^{\aleph_0} = \aleph$ des Kontinuums, und daher gibt es höchstens $\aleph$ verschiedene Gebiete und, als deren Komplemente, höchstens $\aleph$ verschiedene abgeschlossene Mengen. Andererseits gibt es auch mindestens so viele.[1] Sei nämlich A eine isolierte unendliche Menge; die Existenz solcher ist durch § 2 gesichert, und nach II ist A abzählbar,

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}.$$

Zu jedem Punkt a_n läßt sich eine Umgebung $U_{a_n} = V_n$ angeben, die außer a_n keinen Punkt von A enthält. Jeder Menge natürlicher Zahlen $m, n, \ldots$ entspricht dann ein Gebiet

$$G = \mathfrak{S}(V_m, V_n, \ldots),$$

das mit A genau die Punkte $a_m, a_n, \ldots$ gemein hat, und verschiedenen Zahlenmengen entsprechen verschiedene Gebiete. Also:

IV. Die Menge der verschiedenen Gebiete und der verschiedenen abgeschlossenen Mengen ist von der Mächtigkeit des Kontinuums.

Insbesondere gibt es höchstens $\aleph$ endliche Mengen und höchstens $\aleph$ Punkte (vgl. auch S. 272).

Die Gebiete und abgeschlossenen Mengen bilden im euklidischen Raume also nur einen kleinen Teil aller Punktmengen überhaupt, deren Menge ja die Mächtigkeit $2^{\aleph} > \aleph$ hat, und man sieht, daß die Forderung der Abgeschlossenheit oder des Gebietes eine enorme Beschränkung der Allgemeinheit einer Menge ist. Die insichdichten Mengen, in gewissem Sinne das Gegenstück der abgeschlossenen (man vergleiche die definierenden Forderungen $A \geqq A_\beta$ und $A \leqq A_\beta$), bilden hingegen ein System von derselben Mächtigkeit $2^{\aleph}$ wie das System aller Punktmengen; denn beispielsweise ist die Menge R der rationalen Punkte plus einer beliebigen Teilmenge J' der Menge J der irrationalen Punkte eine insichdichte Menge,

[1] Im euklidischen Raume könnten wir uns einfach auf die Menge konzentrischer Kugeln mit beliebigen Radien berufen.

und die J' bilden, da J die Mächtigkeit $\aleph$ hat, ein System von der Mächtigkeit $2^{\aleph}$.

Es gibt, wie der Leser mit Hilfe der Formel $\aleph^{\aleph_0} = \aleph$ zeigen möge, auch nur $\aleph$ abzählbare und $\aleph$ höchstens abzählbare Punktmengen, also $\aleph$ isolierte, separierte, unverdichtete Mengen usw.

Nach Kap. VII, § 2, (6) besteht der Durchschnitt aller Umgebungen von x aus dem Punkt x allein, was eine Formel der Gestalt

$$\{x\} = \mathfrak{D}(U_m, U_n, \ldots) \tag{6}$$

liefert; dieselben Betrachtungen wie zu (5) zeigen wieder, daß E als Punktmenge höchstens von der Mächtigkeit des Kontinuums ist. Natürlich können wir hier nicht allgemein behaupten, daß sie genau von dieser Mächtigkeit sei, da ja jede unendliche Teilmenge von E auf Grund der Relativtheorie wieder ein Raum mit zweitem Abzählbarkeitsaxiom ist.

Ferner gilt der Satz:

V. Die Summe eines beliebigen Systems von Gebieten kann stets durch eine Summe von höchstens abzählbar vielen Gebieten des Systems ersetzt werden.

Den Elementen i einer Menge J seien Gebiete $G_i > 0$ zugeordnet und $S = \underset{i}{\mathfrak{S}} G_i$. Stellt man jedes G_i als Summe von Umgebungen nach (5) dar, so wird nach dem assoziativen Gesetz und mit Beibehaltung nur verschiedener Summanden

$$S = \mathfrak{S}(U_m, U_n, \ldots) = \mathfrak{S}(V_1, V_2, \ldots),$$

wo jedes V_n in mindestens einem G_{i_n} enthalten ist. Danach wird

$$S \leqq \mathfrak{S}(G_{i_1}, G_{i_2}, \ldots) \leqq \underset{i}{\mathfrak{S}} G_i = S,$$

also $S = \mathfrak{S}(G_{i_1}, G_{i_2}, \ldots)$, wo höchstens abzählbar viele Summanden auftreten.

Der Borelsche Satz und seine Umkehrung (Kap. VII, § 4, II III) können hiernach dahin verschärft werden, daß im ersten die Folge von Gebieten durch ein beliebiges System von Gebieten, im zweiten umgekehrt das System durch eine Folge ersetzt werden darf. Sie lauten also dann:

VI (Satz von Borel). Ist eine kompakte abgeschlossene Menge in der Summe eines beliebigen Systems von Gebieten enthalten, so ist sie bereits in einer Summe von endlich vielen Gebieten dieses Systems enthalten.

VII (Umkehrung des Borelschen Satzes). Wenn für jede Folge von Gebieten, in deren Summe die Menge A enthalten ist, A bereits in einer Summe von endlich vielen

Gebieten dieser Folge enthalten ist, so ist A kompakt und abgeschlossen.

Eine letzte Folgerung aus (F) bezieht sich auf relative Dichtigkeit (Kap. VII, § 8). Erinnern wir zuvor daran, daß die durch Entfernungen definierten Umgebungen in einem Raume, in dem eine abzählbare Menge dicht ist, mit einem abzählbaren System von Umgebungen gleichwertig sind (§ 1, Beispiel (4)). Umgekehrt läßt sich, ohne den Begriff Entfernung, aus (F) schließen, daß im Raume eine abzählbare Menge dicht ist, oder allgemeiner:

VIII. Für jede beliebige Menge B gibt es eine zu B dichte, höchstens abzählbare Menge A, insbesondere auch eine solche, die Teilmenge von B (also in B dicht) ist.

Denn sind (von dem trivialen Fall $B=0$ abgesehen) $U_m, U_n, \ldots$ die verschiedenen Umgebungen aller Punkte von B und wählt man aus ihnen irgendwelche Punkte $a_m, a_n, \ldots$, so ist die Menge A der verschiedenen dieser Punkte höchstens abzählbar. Jede Umgebung eines Punktes von B enthält einen Punkt von A, A ist zu B dicht. Insbesondere können die $a_m, a_n, \ldots$ als Punkte von B selbst gewählt werden. In diesem Falle ist $A \leqq B$, $A_\alpha = B_\alpha$, $A_\beta = B_\beta$; also kann jede Menge B_α, d. h. jede abgeschlossene Menge, als Menge der α-Punkte einer höchstens abzählbaren Menge A, und jede Ableitung B_β als Ableitung einer höchstens abzählbaren Menge A dargestellt werden. Daß jede abgeschlossene Menge auch eine Ableitung ist, können wir nur unter einer weiteren Einschränkung beweisen (IX).

Wenn U_m auch Punkte von $E-B$ enthält, so kann a_m unter diesen gewählt werden. Es ist sogar im extremen Fall möglich, A zu B dicht und mit B fremd $(\mathfrak{D}(A, B) = 0)$ zu wählen, nämlich dann und nur dann, wenn jede Umgebung eines Punktes von B auch einen Punkt von $E-B$ enthält, d. h. wenn $E-B$ zu B dicht oder $E-B$ absolut (in E) dicht ist.

In einem metrischen Raume gestattet der Satz VIII noch einige Ergänzungen.

IX. In einem metrischen Raume E mit abzählbarer dichter Teilmenge ist jede abgeschlossene Menge $\leqq E_\beta$ eine Ableitung.

Sei zunächst $A = \{a_1, a_2, \ldots\}$ eine abzählbare Menge $\leqq E_\beta$; wir ordnen jedem Punkte a_n eine sphärische Umgebung V_n mit dem Radius $\frac{1}{n}$ zu, und setzen in jedes V_n eine nach a_n konvergente Menge B_n, was ja möglich ist, weil a_n Häufungspunkt von E sein soll. Für die (abzählbare) Menge $B = \mathfrak{S}(B_1, B_2, \ldots)$ gilt dann zunächst $B_\beta \geqq A$. Andererseits ist aber $B_\beta \leqq A_\alpha$, oder ein nicht zu A_α gehöriger Punkt y gehört auch nicht zu B_β. Denn y hat eine

Umgebung U_y, mit dem Radius 2ϱ, die keinen Punkt von A enthält; für $\frac{1}{n} \leqq \varrho$ hat y von allen Punkten von V_n oder von B_n eine Entfernung $> \varrho$, y ist also kein Häufungspunkt von $\mathfrak{S}(B_n, B_{n+1}, \ldots)$, aber auch keiner von $\mathfrak{S}(B_1, B_2, \ldots, B_{n-1})$, da diese Menge nur die Häufungspunkte $a_1, a_2, \ldots, a_{n-1}$ hat, also ist y kein Häufungspunkt von B. Aus $A \leqq B_\beta \leqq A_\alpha$ folgt durch Bildung der α-Mengen $A_\alpha = B_{\beta\alpha} = B_\beta$; A_α ist die Ableitung von B. Nun kann jede abgeschlossene unendliche Menge als A_α mit abzählbarem A dargestellt werden; und jede abgeschlossene unendliche Menge $\leqq E_\beta$ ist also als B_β darstellbar.

Ist $A = \{a_1, a_2, \ldots, a_m\}$ endlich und B_n eine nach a_n konvergente Menge, so ist ohne weiteres $A = B_\beta$ mit $B = \mathfrak{S}(B_1, B_2, \ldots, B_m)$. Der Satz IX ist damit in allen Fällen bewiesen.

Die Menge B_n kann offenbar jeder Menge entnommen werden, von der a_n Häufungspunkt ist. Ist z. B. H Grenze des Gebietes G, A höchstens abzählbar und in H dicht $(A_\alpha = H)$, so ist jeder Punkt a_n Häufungspunkt des Gebietes G; also kann B_n und B aus G genommen werden. Die Grenze des Gebietes G kann also als Ableitung einer Teilmenge B von G dargestellt werden $(H = B_\beta)$; B ist eine isolierte Menge, da $\mathfrak{D}(B, B_\beta) = \mathfrak{D}(B, H) = 0$. Hiervon wird in der Funktionentheorie häufig Gebrauch gemacht, u. a. zum Beweise des Weierstrass'schen Existenzsatzes, daß es zu einem ebenen (zusammenhängenden) Gebiet G stets eine in G reguläre und über G hinaus nicht fortsetzbare Funktion (eine eindeutige analytische Funktion mit dem Existenzgebiet G) gibt: man zwingt dieser Funktion nämlich die Punkte von B als Nullstellen auf und erreicht dadurch, daß die Punkte der Grenze H singuläre Stellen werden.

X. In jeder kompakten Menge eines metrischen Raumes ist eine höchstens abzählbare Menge dicht.

Ordnen wir jedem Punkt a der kompakten Menge A eine Umgebung U_a mit konstantem Radius ϱ zu, so liegt A in einer endlichen Zahl dieser Umgebungen.[1] Denn andernfalls könnte man eine Folge von Punkten a_n so bestimmen, daß a_2 nicht in U_{a_1} liegt, a_3 nicht in U_{a_1} und U_{a_2} usw.; diese Punkte hätten also paarweise Abstände $\geqq \varrho$ und ihre Menge hätte offenbar keinen Häufungspunkt. — Demgemäß bilden die Mittelpunkte dieser endlich vielen Umgebungen eine endliche Menge $A_\varrho \leqq A$ derart, daß zu jedem Punkt von A ein Punkt von A_ϱ im Abstande $< \varrho$ existiert. Setzt

[1] Das folgt nicht etwa aus dem Borelschen Satz VI, selbst wenn wir A als abgeschlossen annehmen, denn der Satz X soll natürlich unabhängig vom zweiten Abzählbarkeitsaxiom bewiesen werden.

man hierin $\varrho = 1, \frac{1}{2}, \frac{1}{3}, \ldots$, so ist die höchstens abzählbare Menge $\mathfrak{S}(A_1, A_{\frac{1}{2}}, A_{\frac{1}{3}}, \ldots)$ in A dicht.

Danach ist jede kompakte Menge eines metrischen Raumes höchstens von der Mächtigkeit des Kontinuums (S. 272).

§ 4. Punktmengen und Ordnungszahlen.

Bei G. Cantor spielt in der Lehre von den Punktmengen die Theorie der Ordnungszahlen eine wesentliche Rolle, ja sie ist wohl ursprünglich eben zu diesem Zweck ausgebildet worden. E. Lindelöf erkannte, daß der Begriff der Verdichtungspunkte (der übrigens schon bei Cantor vorkommt) die Ordnungszahlen auszuschalten gestattet, und dieser Tendenz folgt auch unsere Darstellung, indem sie den Kern A_k, der bei Cantor als Endglied einer wohlgeordneten Reihe von Mengen erscheint, mit einem Schlage als größte insichdichte Teilmenge von A definiert. Dennoch werden die Ordnungszahlen ihren Platz in der Punktmengentheorie behaupten, aus historischen und sachlichen Gründen. Sie ermöglichen eine feinere Analyse der Struktur gegebener Mengen durch Unterscheidung von Häufungspunkten verschieden hoher Ordnung; eine Analyse, die z. B. bei funktionentheoretischen Untersuchungen wertvolle Dienste geleistet hat (wie bei den Mittag-Lefflerschen Existenzbeweisen für Funktionen mit vorgeschriebenen singulären Stellen), freilich aber dem Mächtigkeitsproblem der Punktmengen keinen Schritt näher kommt als die andere, mehr summarische Behandlung des Unabzählbaren. Natürlich beabsichtigen wir nicht, die Theorie der Punktmengen noch einmal mit Hilfe der Ordnungszahlen aufzubauen, sondern fügen die Relation zwischen beiden in unser System ein. Um uns der Cantorschen Theorie möglichst zu nähern, setzen wir einen Raum voraus, in dem das zweite Abzählbarkeitsaxiom (F) gilt, bemerken aber, daß unsere Betrachtung sich auch ohne diese Annahme durchführen ließe (die niedrigste Mächtigkeit $\aleph_\sigma$ eines Systems von Umgebungen, mit dem das ursprünglich gegebene System der U_x gleichwertig ist, würde dann die Rolle spielen, die in der folgenden Darstellung die Kardinalzahl $\aleph_0$ spielt).

Unter der genannten Voraussetzung haben wir folgenden grundlegenden Satz:

I. Eine auf- oder absteigend wohlgeordnete Menge verschiedener Gebiete oder abgeschlossener Mengen ist höchstens abzählbar.

Der Satz enthält vier Behauptungen.

Erstens sei eine aufsteigend wohlgeordnete Menge von Gebieten

gegeben, d. h. den Ordnungszahlen $0, 1, \ldots, \xi, \ldots$ $(\xi < \lambda)$, deren Menge wir wie früher mit $W(\lambda)$ oder kürzer W bezeichnen, seien Gebiete $G_0, G_1, \ldots, G_\xi, \ldots$ zugeordnet und zwar

$$G_\xi \subset G_\eta \text{ für } \xi < \eta.$$

Die Behauptung lautet, daß W höchstens abzählbar ist (also $\lambda < \omega_1$, λ eine endliche oder zu $Z(\aleph_0)$ gehörige Zahl). Es genügt offenbar, den Beweis für den Fall zu führen, daß W kein letztes Element hat (λ Limeszahl).

Stellen wir nach der Formel

$$G = \mathfrak{S}(U_m, U_n, \ldots)$$

jedes Gebiet als Summe der in ihm enthaltenen Umgebungen dar. Ist $G \subset G'$, so sind alle Summanden von G auch Summanden von G', nicht aber umgekehrt; es gibt also mindestens ein U_n, das in G' enthalten ist, in G aber nicht. Danach ordnen wir jedem Index ξ ein $U_n = V_\xi$ zu, das in $G_{\xi+1}$ enthalten ist, in G_ξ aber nicht. Für $\xi < \eta$ sind dann V_ξ und V_η verschieden; denn es ist $V_\xi \leqq G_{\xi+1} \leqq G_\eta$, während V_η in G_η nicht enthalten ist. Da die Menge aller verschiedenen Umgebungen nach (F) abzählbar, die Menge der V_ξ also höchstens abzählbar ist, so ist die Menge der G_ξ auch nur höchstens abzählbar.

Handelt es sich zweitens um eine absteigend wohlgeordnete Menge von Gebieten, also

$$G_\xi \supset G_\eta \text{ für } \xi < \eta,$$

so sei jetzt umgekehrt V_ξ ein U_n, das in G_ξ enthalten ist, in $G_{\xi+1}$ aber nicht. Für $\xi < \eta$ ist $V_\eta \leqq G_\eta \leqq G_{\xi+1}$, während V_ξ in $G_{\xi+1}$ nicht enthalten ist, also $V_\xi \neq V_\eta$ und weiter wie oben.

Geht man zu den Komplementen über, so entspricht einer auf- oder absteigenden Menge von Gebieten eine ab- oder aufsteigende Menge von abgeschlossenen Mengen, womit der Satz I vollständig bewiesen ist.

Der Satz gilt auch für Relativgebiete und relativ abgeschlossene Teilmengen einer Menge M; denn entwickelt man die Relativtheorie (Kap. VII, § 6) für M als umfassenden Raum, so sind die in Frage kommenden Relativumgebungen die Mengen $\mathfrak{D}(M, U_1)$, $\mathfrak{D}(M, U_2), \ldots$ und die verschiedenen unter ihnen bilden ein höchstens abzählbares System. Also:

II. Eine auf- oder absteigend wohlgeordnete Menge verschiedener Relativgebiete von M oder verschiedener in M abgeschlossener Mengen ist höchstens abzählbar.

Bezeichnen wir die fraglichen Mengen mit A_ξ und lassen jetzt nur die Voraussetzung ihrer Verschiedenheit fallen, ersetzen also die

Relationen $A_\xi < A_\eta$, $A_\xi > A_\eta$ resp. durch $A_\xi \leqq A_\eta$, $A_\xi \geqq A_\eta$, so folgt, daß im Falle einer unabzählbaren Menge $W(\lambda)$ von Indices (also für $\lambda \geqq \omega_1$) höchstens abzählbar viele A_ξ von einander verschieden sein können. Es gibt also eine Ordnungszahl $\eta < \omega_1$ derart, daß

$$A_\eta = A_{\eta+1} = \ldots = A_\zeta = \ldots \qquad (\zeta > \eta),$$

d. h. es wird nach einer höchstens abzählbaren Menge von Schritten eine letzte Menge A_η in dem Sinne erreicht, daß alle folgenden mit ihr übereinstimmen.

Hiervon machen wir folgende Anwendung. Vermöge irgend eines Gesetzes sei jeder Menge M eine in M abgeschlossene Menge $\varphi(M)$ zugeordnet. Für alle endlichen oder abzählbaren Ordnungszahlen $\xi < \omega_1$ definieren wir dann durch Induktion eine Menge A_ξ folgendermaßen:

A_0 sei eine beliebig gegebene Menge A, und für $\eta > 0$ sei

(1) $$A_\eta = \underset{\xi}{\mathfrak{D}} \varphi(A_\xi) \qquad (\xi < \eta).$$

Ähnlich wie bei den durch Iteration definierten Normalfunktionen (S. 116) erkennt man sofort: für $\xi < \eta$ ist

$$\varphi(A_\eta) \leqq A_\eta \leqq \varphi(A_\xi) \leqq A_\xi,$$

die A_ξ und $\varphi(A_\xi)$ bilden also eine absteigend wohlgeordnete Reihe von (nicht notwendig verschiedenen) Mengen. Daraus folgt, daß

(2) $$A_{\xi+1} = \varphi(A_\xi),$$

also

$$A_\eta = \underset{\xi}{\mathfrak{D}} A_{\xi+1} \qquad (\xi < \eta)$$

und für eine Limeszahl η auch

(3) $$A_\eta = \underset{\xi}{\mathfrak{D}} A_\xi \qquad (\xi < \eta)$$

ist; die ersten dieser Mengen sind

$$A_0 = A,\ A_1 = \varphi(A_0),\ A_2 = \varphi(A_1), \ldots$$
$$A_\omega = \mathfrak{D}(A_0, A_1, A_2, \ldots),\ A_{\omega+1} = \varphi(A_\omega), \ldots.$$

Durch Induktion folgt aus (1), daß alle diese Mengen in A abgeschlossen sind; denn ist A_ξ in A abgeschlossen, so ist $\varphi(A_\xi)$ in A_ξ, also auch in A abgeschlossen, und A_η als Durchschnitt solcher Mengen ebenfalls.

Nach II wird nun eine letzte (kleinste) Menge A_η erreicht, mit der alle folgenden übereinstimmen, und zwar geschieht dies bereits, wenn $A_\eta = A_{\eta+1}$; wir verstehen unter η die kleinste Zahl[1], für die

---

[1] Wenn $A_\eta = 0$ und A eine abgeschlossene kompakte Menge ist, so ist η keine Limeszahl. Das folgt aus dem Cantorschen Durchschnittssatz (Kap. VII, § 4, I) unter Berücksichtigung, daß η als Limeszahl mit ω konfinal wäre.

diese Gleichung besteht, so daß, für $\xi < \eta$, $A_\xi > A_{\xi+1}$ ist. Setzen wir noch

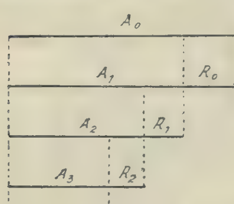


Fig. 9.

(4) $$R_\xi = A_\xi - A_{\xi+1} = A_\xi - \varphi(A_\xi),$$

so gilt die Zerlegung

(5) $$A = \sum_\xi R_\xi + A_\eta \qquad (\xi < \eta).$$

Denn ein Punkt von A gehört entweder zu A_η oder nicht; im letzten Fall sei A_ζ $(\zeta \leqq \eta)$ die Menge mit kleinstem Index, der er nicht angehört. Dann ist $\zeta > 0$ und nach (3) keine Limeszahl; setzen wir also $\zeta = \xi + 1$ $(\xi < \eta)$, so gehört der fragliche Punkt noch zu A_ξ, aber nicht mehr zu $A_{\xi+1}$, sondern also zu R_ξ.

Erstes Beispiel: Kohärenzen. Es sei

(A) $$\varphi(M) = \mathfrak{D}(M, M_\beta) = M_h$$

die Menge der zu M gehörigen Häufungspunkte von M; das ist, da M_β abgeschlossen ist, eine in M abgeschlossene Menge, die G. Cantor die (erste) Kohärenz von M genannt hat. Die Differenz

$$M - \varphi(M) = M_j = M_\alpha - M_\beta$$

ist die Menge der isolierten Punkte von M, also höchstens abzählbar und übrigens eine Differenz abgeschlossener Mengen (die Adhärenz von M). Dieser Prozeß führt also, jedesmal durch Abspaltung der isolierten Punkte, zu den Kohärenzen

$$A_0 = A, \; A_1 = A_h, \; A_2 = A_{hh}, \; A_3 = A_{hhh}, \ldots,$$

dann zu

$$A_\omega = \mathfrak{D}(A, A_h, A_{hh}, \ldots), \; A_{\omega+1} = A_{\omega h}, \ldots$$

und so fort bis zur letzten Kohärenz A_η, die mit $A_{\eta+1} = A_{\eta h}$ übereinstimmt, also insichdicht (aber eventuell Null) ist. Diese letzte Kohärenz ist der Kern von A, d. h. die größte insichdichte Teilmenge von A oder, wie wir sie erklärt hatten, die Summe aller insichdichten Teilmengen von A. Denn eine insichdichte Teilmenge von M ist auch Teilmenge von M_h; eine insichdichte Teilmenge von A gehört also, wie sofort folgt, allen Kohärenzen an, und A_k ist demgemäß in A_η, aber A_η als insichdichte Menge auch in A_k enthalten. In der Zerlegung (5) ist also A_η der Kern, der andere Bestandteil oder die Summe der Adhärenzen

$$\sum R_\xi = \sum A_{\xi j} = A_s$$

der separierte Bestandteil von A, und wir sehen, wie diese von

uns mit einem Schlage[1] definierten Mengen auch durch sukzessive (aber ins Unendliche fortgesetzte) Abspaltung der isolierten Punkte erhalten werden können. Aus der letzten Formel schloß auch G. Cantor, was wir schon anderweitig wissen, daß A_s höchstens abzählbar ist, da A_s als Summe von höchstens abzählbar vielen isolierten, also höchstens abzählbaren Mengen R_ξ erscheint.

Ist die Menge A speziell abgeschlossen, so gehen die Kohärenzen in die Ableitungen

$$A_0 = A, \; A_1 = A_\beta, \; A_2 = A_{\beta\beta}, \dots$$

über und man gelangt zu einer letzten Ableitung A_η, die den (in diesem Falle perfekten) Kern A_k darstellt.

Betrachten wir von einer beliebigen Menge A und zugleich von $B = A_\alpha$ die Kohärenzen

$$A_0 = A, \; A_1 = A_h, \; A_2 = A_{hh}, \dots,$$
$$B_0 = B, \; B_1 = B_h, \; B_2 = B_{hh}, \dots;$$

die letzteren sind die Ableitungen von B, also wegen $B_\beta = A_{\alpha\beta} = A_\beta$ auch von A, d. h.

$$B_1 = A_\beta, \; B_2 = A_{\beta\beta}, \dots .$$

Die letzte Kohärenz von A ist der Kern A_k, die letzte Ableitung von A der Kern B_k, wobei $A_k \leqq B_k$. Das Verschwinden der letzten Ableitung zieht das der letzten Kohärenz nach sich, aber nicht umgekehrt.

Ist z. B. (wie S. 255) A eine isolierte ebene Menge, deren Ableitung die Abszissenachse oder sonst eine perfekte Menge ist, so ist schon die erste Kohärenz $A_h = 0$, während alle Ableitungen $A_\beta = A_{\beta\beta} = \dots$ von Null verschieden sind.

Daß es wirklich Mengen gibt, deren letzte Kohärenz A_η erst nach einer unendlichen Reihe von Schritten und sogar mit beliebig großem Index η ($< \omega_1$) erreicht wird, lehren die wohlgeordneten linearen Mengen (Mengen reeller Zahlen). Eine Menge reeller Zahlen ist, der Größe nach, eine geordnete Menge; nach Kap. IV, § 7, I gibt es in der dichten Menge aller reellen Zahlen abzählbare Mengen von beliebigem, insbesondere wohlgeordnetem Typus. Umgekehrt ist übrigens eine wohlgeordnete Menge $A = \{a_1, a_2, \dots\}$ reeller Zahlen höchstens abzählbar, da die zugehörigen Halbgeraden ($x \geqq a_1$, $x \geqq a_2$, ...) eine absteigende Reihe abgeschlossener Mengen bilden. Beschränken wir uns zur Vermeidung von Ausnahmefällen auf Mengen $A = \{a_1, a_2, \dots, a_\lambda\}$ mit letztem Element; A ist mit der

---

[1] Dafür haben wir, wie eingangs bemerkt wurde, hier eine feinere Zergliederung der Menge A und ihres separierten Bestandteils A_s; man kann die Mengen nach dem Index η der letzten Kohärenz klassifizieren u. dgl.

Menge $W = \{1, 2, \ldots, \lambda\}$ ähnlich.[1] Dann ist A_β mit der Menge der in W enthaltenen Limeszahlen ähnlich. Denn ist x $(\leqq a_\lambda)$ eine Häufungszahl von A, und η der kleinste Index, für den $x \leqq a_\eta$, also $a_\xi < x \leqq a_\eta$ für $\xi < \eta$, so muß η Limeszahl und x mit der kleinsten Zahl x_η, die alle a_ξ übertrifft, identisch sein. Die Menge der in W enthaltenen Limeszahlen ist nun $W_1 = \{\omega 1, \omega 2, \ldots, \omega\mu\}$, wobei μ die durch $\omega\mu \leqq \lambda < \omega(\mu + 1)$ bestimmte Zahl ist (Kap. V, S. 109), und $A_\beta = \{x_{\omega 1}, x_{\omega 2}, \ldots, x_{\omega\mu}\}$ ist mit W_1 oder $W' = \{1, 2, \ldots, \mu\}$ ähnlich; für $\mu = 0$ verschwinden alle diese Mengen. Umgekehrt kann man hiernach, bei vorgeschriebenem μ, eine mit W' ähnliche abgeschlossene lineare Menge A_β konstruieren (indem man z. B. $\lambda = \omega\mu$ wählt); demgemäß wollen wir sogleich A als abgeschlossen annehmen, und $A_\beta = \{a_{\omega 1}, a_{\omega 2}, \ldots, a_{\omega\mu}\}$ ist dann die Menge derjenigen a, deren Indices Limeszahlen sind oder die Menge W_1 durchlaufen. Schreibt man dafür $A_\beta = \{b_1, b_2, \ldots, b_\mu\}$, so ist $A_{\beta\beta} = \{b_{\omega 1}, b_{\omega 2}, \ldots, b_{\omega\nu}\}$ die Menge derjenigen b, deren Indices Limeszahlen sind, oder derjenigen a, deren Indices Vielfache von ω^2 sind, d. h. die Menge $W_2 = \{\omega^2 1, \omega^2 2, \ldots, \omega^2 \nu\}$ durchlaufen; hier ist ν durch $\omega^2\nu \leqq \lambda < \omega^2(\nu + 1)$ bestimmt. Allgemein behaupten wir: die ξ^{te} Ableitung A_ξ ist die Menge derjenigen a, deren Indices Vielfache von ω^ξ sind, also die Menge

$$W_\xi = \{\omega^\xi 1, \omega^\xi 2, \ldots, \omega^\xi \lambda_\xi\}$$

durchlaufen, wo $\omega^\xi \lambda_\xi \leqq \lambda < \omega^\xi(\lambda_\xi + 1)$ ist: für $\lambda_\xi = 0$ verschwinden W_ξ und A_ξ. Diese Behauptung ist durch Rekursion leicht zu beweisen: ihre Übertragbarkeit von ξ auf $\xi + 1$ liegt auf der Hand, und für eine Limeszahl η ist $W_\eta = \mathfrak{D}_\xi W_\xi$, d. h. eine Zahl, die gleichzeitig Vielfaches aller ω^ξ (für $\xi < \eta$) ist, ist auch Vielfaches von ω^η (ist $\alpha = \omega^\eta\beta + \gamma$, $\gamma < \omega^\eta$, so kann nicht stets $\gamma \geqq \omega^\xi$ sein; ist also $\gamma < \omega^\xi$ für irgend ein ξ, so ist $\alpha = \omega^\xi \cdot \omega^{-\xi+\eta}\beta + \gamma$ durch ω^ξ nur teilbar für $\gamma = 0$, d. h. α ist, wenn durch alle ω^ξ, auch durch ω^η teilbar).

Danach hat z. B. für $\lambda \geqq \omega^\xi$ die Menge A, vom Typus $\lambda + 1$, eine ξ^{te} Ableitung, die sicher nicht verschwindet; andererseits ist für $\omega^\eta > \lambda$ die η^{te} Ableitung Null, und die letzte Ableitung Null wird also erst für einen Index $> \xi$ erreicht. Für $\lambda = \omega^\xi$ besteht A_ξ aus a_λ allein, und die nächste Ableitung $A_{\xi+1}$ ist Null.

Zweites Beispiel: Residuen. Es sei

$$\psi(M) = M_\alpha - M$$

---

[1] Für $\lambda = 0$ verstehen wir unter W und A die Nullmenge, wie im Folgenden zu beachten ist.

die Menge der nicht zu M gehörigen Häufungspunkte von M (der Rand des Komplements $E - M$) und

(B) $$\varphi(M) = \psi(\psi(M))$$

die durch zweimalige Anwendung von ψ entstehende Menge. Während $\psi(M)$ zu M fremd ist, ist $\varphi(M)$ eine in M abgeschlossene Menge; denn setzt man

$$\psi(M) = N, \ \varphi(M) = M',$$

also

$$M_\alpha = M + N, \ N_\alpha = M' + N,$$

so ist $N \leqq M_\alpha$, also $N_\alpha \leqq M_\alpha$, folglich $M' \leqq M$ und

$$M' = \mathfrak{D}(M, N_\alpha).$$

Wir bemerken noch, daß wieder wie bei dem vorigen Prozeß

$$M - \varphi(M) = M_\alpha - N_\alpha$$

eine Differenz abgeschlossener Mengen ist, die aber nicht mehr endlich oder abzählbar zu sein braucht.

Nach Definition ist $\psi(M) \leqq M_\beta$, also $\varphi(M) \leqq M_{\beta\beta}$ und jedenfalls $\varphi(M) \leqq M_h$; auch dieser Prozeß befreit M also von den isolierten Punkten. Aber er ist im allgemeinen viel energischer als die Kohärenzbildung, indem er z. B. auch alle inneren Punkte abspaltet. Denn $\psi(M)$ ist in der Grenze M_g, $\varphi(M)$ in deren Ableitung enthalten; da M_g abgeschlossen ist, so ist auch $\varphi(M) \leqq M_g$, also $\varphi(M) \leqq M_r$. Eine abgeschlossene Menge oder ein Gebiet wird durch diesen Prozeß sofort auf Null reduziert. Wir wollen $\varphi(M)$ das (erste) Residuum von M nennen.

Definieren wir, wie oben, von einer beliebigen Menge A ausgehend die Reihe der Residuen

$$A_0 = A, \ A_1 = \varphi(A_0), \ A_2 = \varphi(A_1), \ldots,$$

so gelangen wir wieder für einen Index $\eta < \omega_1$ zu einem letzten Residuum A_η, für welches also $A_\eta = \varphi(A_\eta)$. Diese Menge, die wir in ihrer Abhängigkeit von A mit A_l bezeichnen wollen, ist nach dem Gesagten insichdicht (also $A_l \leqq A_k$) und eine Randmenge. Wenn sie Null ist, wollen wir A reduzibel[1] nennen. Nach (5) setzt sich A aus A_l und einer Summe von höchstens abzählbar vielen Differenzen (R_ξ) abgeschlossener Mengen zusammen; eine reduzible Menge ist selbst eine solche Summe.

Zu den reduziblen Mengen gehören die separierten, aber noch viele andere, so die abgeschlossenen Mengen und die Gebiete. Allgemein sind alle Mengen, die aus abgeschlossenen Mengen

---

[1] G. Cantor bezeichnet, in viel engerem Sinne, als „reduktible" Menge den separierten Bestandteil einer abgeschlossenen Menge.

durch wiederholte Differenzbildung entstehen, d. h. die Mengen des kleinsten Körpers über dem Ring der abgeschlossenen Mengen (Kap. I, § 7) reduzibel, und zwar wird hier das letzte Residuum $A_\eta = 0$ schon nach einer endlichen Zahl von Schritten erreicht ($\eta < \omega$). In der Tat läßt sich nach der zitierten Stelle eine solche Menge A als Differenzenkette

$$A = P_1 - P_2 + P_3 - P_4 + \dots$$

mit endlicher Gliederzahl n darstellen, wo $P_1 \geqq P_2 \geqq \dots \geqq P_n$ abgeschlossene Mengen sind. Schneiden wir mit A_α, so entsteht nach dem distributiven Gesetz

$$A = \mathfrak{D}(A, A_\alpha) = Q_1 - Q_2 + Q_3 - Q_4 + \dots,$$

wobei $Q_m = \mathfrak{D}(P_m, A_\alpha)$ wieder abgeschlossen ist. Überdies ist $A \leqq P_1$, $A_\alpha \leqq P_1$, also $Q_1 = A_\alpha$, und nach der Bedeutung der Differenzenketten

$$\psi(A) = A_\alpha - A = Q_2 - Q_3 + Q_4 - \dots$$

eine Differenzenkette von $n-1$ Gliedern (oder weniger, falls etwaige verschwindende weggelassen werden). Der nächste Schritt gibt $\varphi(A)$ als Differenzenkette von höchstens $n-2$ Gliedern, sodaß in der Reihe der Residuen nach endlicher Zahl von Wiederholungen die Null auftritt. Umgekehrt, eine reduzible Menge mit endlichem Index des letzten Residuums ist eine Summe von endlich vielen Mengen R_ξ, also dem genannten Körper angehörig.

Um noch weitere reduzible Mengen festzustellen, bemerken wir: ist M in A abgeschlossen, so ist auch $\psi(M)$ in $\psi(A)$ und M_ι in A_ι abgeschlossen. Denn es ist

$$\psi(M) = M_\alpha - M = \mathfrak{D}(M_\alpha, A_\alpha - A)$$

in $\psi(A)$ abgeschlossen, weiter $\psi\psi(M) = \varphi(M)$ in $\psi\psi(A) = \varphi(A)$ abgeschlossen, also, wie durch Induktion unmittelbar folgt, jedes Residuum M_ξ in dem entsprechenden Residuum A_ξ abgeschlossen. Demnach ist jede in einer reduziblen Menge abgeschlossene Menge selbst reduzibel. Für beliebige Teilmengen gilt dies nicht; sonst wäre, da der Raum E reduzibel ist, überhaupt jede Menge reduzibel, während z. B. die Menge R der rationalen Punkte des euklidischen Raumes mit $\varphi(R)$ und daher mit R_ι übereinstimmt.

Ist M in A abgeschlossen und $M = \varphi(M)$, so ist $M = M_\iota$ in A_ι abgeschlossen; demnach ist das letzte Residuum A_ι die größte in A abgeschlossene Teilmenge, für die $M = \varphi(M)$.

Ist $A + B = C$ abgeschlossen und M in A abgeschlossen, so ist $\psi(M) = \mathfrak{D}(M_\alpha, B)$ in B abgeschlossen. Also ist $\psi(A_\iota)$ in B abgeschlossen, und da $\varphi\psi(A_\iota) = \psi\psi\psi(A_\iota) = \psi\varphi(A_\iota) = \psi(A_\iota)$, so ist $\psi(A_\iota)$ auch in B_ι und ebenso $\psi(B_\iota)$ in A_ι abgeschlossen; andererseits aber

ist auch $\psi\psi(A_l)$ in $\psi(B_l)$ oder A_l in $\psi(B_l)$ abgeschlossen, also endlich

$$A_l = \psi(B_l), \quad B_l = \psi(A_l).$$

Demnach ist das Komplement einer reduziblen Menge in einer abgeschlossenen Menge selbst reduzibel. Z. B. sind die Mengen A, $E-A$ und $\psi(A)$ gleichzeitig reduzibel oder nicht.

Wir deuten noch an, wie man das letzte Residuum analog zur letzten Kohärenz ohne Ordnungszahlen einführen kann. Legt man eine beliebige, aber festgewählte Menge A und ihr Komplement $B = E - A$ zugrunde und definiert die Funktion $\chi(M)$ durch

$$N = \mathfrak{D}(M_\alpha, B), \quad \chi(M) = \mathfrak{D}(N_\alpha, A),$$

so ist $\chi(M)$ mit M kogredient (d. h. für $M_1 \leqq M$ auch $\chi(M_1) \leqq \chi(M)$, was für $\varphi(M)$ nicht zutrifft). Daraus geht unmittelbar hervor, daß eine Summe beliebig vieler Mengen M, für die $M \leqq \chi(M)$, wieder dieselbe Eigenschaft hat, genau wie sich die Insichdichtheit von beliebig vielen Mengen auf ihre Summe überträgt, und daß man daraufhin die größte Teilmenge M von A von dieser Eigenschaft definieren kann. Das ist aber das letzte Residuum, wie sich leicht zeigen läßt, wenn man beachtet, daß, für jede in A abgeschlossene Menge M, $\chi(M) = \varphi(M)$ ist.

Um auch Mengen mit beliebig hohem Index des letzten Residuums zu bilden, verstehen wir unter M eine separierte abgeschlossene Menge (z. B. eine wohlgeordnete abgeschlossene Menge reeller Zahlen, S. 280), unter M_ξ ihre Ableitungen und unter

$$\begin{aligned} A &= (M - M_1) + (M_2 - M_3) + \dots + (M_\omega - M_{\omega+1}) + \dots \\ &= \sum_\xi (M_{2\xi} - M_{2\xi+1}) \end{aligned}$$

die Summe ihrer Adhärenzen gerader Ordnung. Setzen wir

$$A = M - B = (M - M_1) + A';$$

da $M - M_1 \leqq A \leqq M$ und die Menge $M - M_1 = M_i$ der isolierten Punkte von M in M dicht ist (S. 255), so ist auch A in M dicht und $A_\alpha = M$, folglich

$$\psi(A) = M - A = B = M_1 - A'.$$

Weiter ist $M - M_1 \leqq A \leqq (M - M_1) + M_2$, also $M_1 - M_2 \leqq B \leqq M_1$ und auf Grund derselben Schlußweise

$$\psi(B) = M_1 - B = A'$$

oder

$$\varphi(A) = (M_2 - M_3) + (M_4 - M_5) + \dots.$$

Hieraus folgt durch Induktion, daß das ξ^{te} Residuum von A

$$A_\xi = \sum_\eta (M_{2\eta} - M_{2\eta+1}) \qquad (\eta \geqq \xi)$$

und von Null verschieden ist, sobald die $2\xi^{\text{te}}$ Adhärenz von M nicht verschwindet, was sich ja für beliebig großes ξ erreichen ließ (S. 280); das letzte Residuum A_η ist Null. Weiteres über reduzible Mengen s. Anhang.

Drittes Beispiel: Komponenten. Unter $\varphi(M)$ verstehen wir jetzt, wenn M eine nichtverschwindende, unzusammenhängende Menge ist, eine nichtverschwindende in M abgeschlossene Menge, deren Komplement $M - \varphi(M)$ ebenfalls in M abgeschlossen und von Null verschieden ist; ist M zusammenhängend oder Null, so sei $\varphi(M) = M$. Die Reihe

$$A_0 = A,\ A_1 = \varphi(A_0),\ A_2 = \varphi(A_1), \ldots$$

führt dann zu einer letzten Menge $A_\eta = \varphi(A_\eta)$, die entweder zusammenhängend oder Null ist. Ist sie zusammenhängend, so ist sie eine Komponente von A; denn da eine Komponente von M entweder $\leqq \varphi(M)$ oder $\leqq M - \varphi(M)$ ist, so muß diejenige Komponente C von A, die mit A_η, also mit sämtlichen A_ξ Punkte gemeinsam hat, in allen A_ξ enthalten sein, woraus gleichzeitig $C \leqq A_\eta$ und $C \geqq A_\eta$, also $C = A_\eta$ folgt. Der Fall $A_\eta = 0$ kann (falls A nicht verschwindet) nur für eine Limeszahl η eintreten und läßt sich durch die Vorschrift verhindern, daß jedes A_ξ einen gegebenen Punkt a von A enthalten soll; er tritt sicher, bei beliebiger Wahl unter den möglichen Mengen A_ξ, nicht ein, wenn die Menge A abgeschlossen, kompakt und von Null verschieden ist.

§ 5. Mengen mit Raumcharakter.

Bevor wir unsere Betrachtung weiter spezialisieren, ist es vielleicht an dieser Stelle angebracht, einige Beispiele von Mengen zu geben, deren „Punkte" und Umgebungen von denen des euklidischen Raumes bisweilen recht verschieden sind und die trotzdem alle unsere bisherigen Voraussetzungen oder einen Teil von ihnen erfüllen. Für solche Mengen gelten also die entsprechenden Sätze unserer bisherigen Theorie. Die Deutung dieser Sätze bedarf in jedem Einzelfall spezieller Untersuchung, auf die wir größtenteils verzichten wollen; man hat z. B. festzustellen, was die Konvergenz einer Folge ($\lim a_n = x$ oder in einem metrischen Raume $\lim \overline{xa_n} = 0$) besagt, unter welchen Bedingungen eine Menge kompakt ist usw.

I. Teilweise metrische Räume.

Wir wollen E einen teilweise metrischen Raum nennen (in Analogie zu den teilweise geordneten Mengen), wenn nur einem Teil seiner Punktpaare, nicht notwendig allen, Entfernungen zugeordnet sind, die aber, soweit sie existieren, die Entfernungsaxiome (S. 211) erfüllen. Dem Dreiecksaxiom ist hierbei zweckmäßig folgende Fassung zu

geben: wenn die Entfernungen $\overline{xy}$ und $\overline{yz}$ definiert sind, so ist auch $\overline{xz}$ definiert und $\leqq \overline{xy} + \overline{yz}$. Auf Grund dieser Festsetzung haben zwei Punkte, die von einem dritten eine Entfernung haben, auch untereinander eine Entfernung: die Menge $E(x)$ der Punkte, die von x eine Entfernung haben, ist ein metrischer Raum, und E zerfällt in eine Summe paarweise fremder metrischer Räume $E(x) + E(y) + \ldots$. Definieren wir nun wieder eine sphärische Umgebung U_x mit dem Radius ϱ als Menge der Punkte y mit $\overline{xy} < \varrho$, so sieht man unmittelbar, daß auch jetzt die Umgebungsaxiome erfüllt bleiben (bei (D) ist, falls x, y eine Entfernung haben, der Beweis wie damals S. 214; haben x, y keine Entfernung, so ist stets $\mathfrak{D}(U_x, U_y) = 0$). Ebenso bleiben die Folgerungen aus dem ersten und, falls in E eine abzählbare Menge R dicht ist, auch aus dem zweiten Abzählbarkeitsaxiom richtig. Die Konvergenz einer Punktfolge $\lim a_n = x$ hat jetzt die Bedeutung, daß für fast alle n die Entfernung $\overline{xa_n}$ definiert ist und mit $\frac{1}{n}$ nach Null konvergiert.

II. Adjunktion uneigentlicher Punkte.

Wir fügen einem metrischen Raume E mit sphärischen Umgebungen U_x ein weiteres Element hinzu, das wir mit ∞ bezeichnen, wodurch die Menge $E_\infty = E + \{\infty\}$ entsteht. In dieser Menge E_∞ sollen die U_x nach wie vor die Umgebungen von Punkten x der Menge E bleiben; unter einer Umgebung U_∞ aber verstehen wir das Äußere einer Kugel, nämlich die Menge der Punkte y mit $\overline{ay} > \varrho$ (wo a ein fester Punkt von E und ϱ eine positive Zahl ist), nebst dem Punkte ∞ selbst. In dieser Weise ergänzt ja die Funktionentheorie die gewöhnliche Ebene durch Hinzufügung des „unendlich fernen Punktes“. Man sieht, daß die Umgebungsaxiome und das erste Abzählbarkeitsaxiom erfüllt bleiben. Wenn in E eine abzählbare Menge R dicht ist, so sind diese Umgebungen wieder mit einem System solcher gleichwertig, das auch das zweite Abzählbarkeitsaxiom erfüllt, nämlich mit dem System der Umgebungen von Punkten der Menge $R + \{\infty\}$ mit rationalen Radien.

III. Nichtarchimedische Entfernungen.

In einer Menge E seien Entfernungen $\overline{xy}$ definiert, die aber nicht mehr reelle Zahlen, sondern Elemente eines nichtarchimedischen Größensystems (Kap. VI, § 11) sind, von dem wir voraussetzen wollen, daß zu jeder Größe α auch der n^{te} Teil $\frac{1}{n}\alpha$ im System vorhanden sei. Die Größen $\overline{xy}$ sollen natürlich $\geqq 0$ sein und die Entfernungsaxiome erfüllen. Aus der größtenteils wörtlichen Wiederholung der Überlegung in Kap. VII, § 1 erkennt man, daß die durch

$\overline{xy} < \varrho$ mit positivem ϱ definierten Umgebungen U_x wieder die Umgebungsaxiome erfüllen (die Schlüsse beruhen darauf, daß es zwischen zwei Größen $\alpha < \beta$ des Systems eine dritte, z. B. $\frac{1}{2}(\alpha + \beta)$ gibt). Den einfachsten Fall, den linearen Mengen entsprechend, erhalten wir, wenn wir unter E das nichtarchimedische Größensystem selbst verstehen und als Entfernung $\overline{xy}$ den Betrag $|x - y|$ definieren, wobei natürlich $|\alpha|$ diejenige von den beiden Größen $\pm\alpha$ bedeutet, die $\geqq 0$ ist. Die Umgebung U_x mit dem Radius ϱ besteht dann aus denjenigen Punkten (Größen) y, für die $x - \varrho < y < x + \varrho$ ist.[1]

Die Menge der positiven Größen des Systems ist, da sie nach Annahme kein kleinstes Element hat, mit dem Inversen einer regulären Anfangszahl ω_σ koinitial, d. h. es gibt eine Menge positiver Größen

$$\varrho_0 > \varrho_1 > \varrho_2 > \cdots > \varrho_\xi > \cdots \qquad (\xi < \omega_\sigma),$$

zu denen keine noch kleinere positive Größe existiert.[2] Offenbar gilt das erste Abzählbarkeitsaxiom dann und nur dann, wenn diese Anfangszahl $\omega_\sigma = \omega$ ist.

Nehmen wir für das Größensystem E den einfachsten Fall: es bestehe aus sämtlichen, lexikographisch geordneten Komplexen reeller Zahlen

$$x = (x_0, x_1, \ldots, x_\lambda, \ldots) \qquad (\lambda < \mu)$$

mit dem wohlgeordneten Argument $W(\mu)$, d. h. der Menge der Ordnungszahlen $< \mu$. Die Ordnungszahl μ ist > 1 vorauszusetzen, da wir für $\mu = 1$ auf die Menge der reellen Zahlen zurückkämen. Je nachdem μ einen Vorgänger $\mu - 1$ hat oder Limeszahl ist, ist die Menge der positiven Komplexe mit ω^* oder μ^* koinitial; im ersten Fall nämlich mit der Menge der Komplexe, die an der letzten Stelle $\mu - 1$ eine der Zahlen $1, \frac{1}{2}, \frac{1}{3}, \ldots$ und sonst lauter Nullen haben, im zweiten Fall mit der Menge derer, die nur an einer einzigen Stelle eine Eins und sonst lauter Nullen haben. Das erste Abzählbarkeitsaxiom gilt also dann und nur dann, wenn μ einen Vorgänger hat oder mit ω konfinal ist. — Das zweite Abzählbarkeitsaxiom ist keinesfalls gültig, weil eine in E dichte Menge A niemals abzählbar ist. Denn A muß, für jede reelle Zahl x_0, einen zwischen $(x_0, -1, \ldots)$ und $(x_0, +1, \ldots)$ liegenden, also mit x_0 beginnenden Komplex enthalten und demnach mindestens von der Mächtigkeit des Kontinuums sein.

---

[1] Schließlich kann man sich hier von Metrik ganz befreien und in einer geordneten Menge E (ohne erstes und letztes Element) unter den U_x die den Punkt x enthaltenden Mittelstrecken verstehen (S. 214).

[2] Anders ausgedrückt: die Größe 0, und damit jede Größe des Systems, ist ω_σ^*-Limes und zugleich ω_σ-Limes.

Der Leser diskutiere den Fall $\mu = 2$, wo die Komplexe $x = (x_0, x_1)$ wieder als Punkte der Ebene gedeutet werden können, aber die Umgebungen eine ganz andere Form als gewöhnlich haben: es sind Parallelstreifen mit je einer Hälfte der einschließenden Geraden oder Strecken parallel der Achse $x_0 = 0$, je nachdem der „Radius“ $\varrho = (\varrho_0, \varrho_1)$ positives oder verschwindendes ϱ_0 (im letzteren Falle positives ϱ_1) hat. Eine Gerade $x_1 = \text{const}$ hat keinen Häufungspunkt, wie im Beispiel § 1, (2).

IV. Der Raum mit unendlich vielen Dimensionen.

Die Elemente der Menge E seien reelle Zahlenfolgen

$$x = (x_1, x_2, \ldots, x_n, \ldots),$$

für welche die Summe

$$x_1^2 + x_2^2 + \cdots + x_n^2 + \cdots$$

konvergiert. Für zwei solche Folgen x, y setzen wir

$$f_n(x, y) = (x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2$$

und im Falle der Konvergenz

$$f(x, y) = \lim f_n(x, y) = (x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots.$$

Es ist

$$\sqrt{f_n(x, y)} \leqq \sqrt{f_n(x, z)} + \sqrt{f_n(y, z)},$$

wenn die Wurzeln $\geqq 0$ genommen werden; das ist nichts anderes als die Dreiecksrelation im euklidischen Raume E_n. Insbesondere ist, wenn wir für z die Nullfolge setzen,

$$\sqrt{f_n(x, y)} \leqq \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} + \sqrt{y_1^2 + y_2^2 + \cdots + y_n^2},$$

und aus der angenommenen Konvergenz der Quadratsummen folgt

$$\sqrt{f_n(x, y)} \leqq \sqrt{x_1^2 + x_2^2 + \cdots} + \sqrt{y_1^2 + y_2^2 + \cdots},$$

daraus aber die Konvergenz der (aufsteigenden, beschränkten) Folge $f_n(x, y)$. $f(x, y)$ existiert also stets und aus der Dreiecksrelation für den E_n folgt durch Grenzübergang

$$\sqrt{f(x, y)} \leqq \sqrt{f(x, z)} + \sqrt{f(y, z)}.$$

Definieren wir also die Entfernung durch $\overline{xy} = \sqrt{f(x, y)}$, so sind die Entfernungsaxiome erfüllt, und E ist ein metrischer Raum, den man nach Analogie als einen euklidischen Raum mit abzählbar vielen Dimensionen bezeichnen kann; er wird auch der Hilbertsche Raum genannt. Die sphärischen Umgebungen seiner Punkte erfüllen also die Umgebungsaxiome.

In ihm ist eine abzählbare Menge dicht, nämlich die Menge R der Punkte

$$r = (r_1, r_2, \ldots, r_n, 0, 0, 0, \ldots),$$

die rationale und schließlich verschwindende Koordinaten haben.[1] Denn ist x ein beliebiger Punkt und ϱ eine positive Zahl, so kann man n so wählen, daß

$$x_{n+1}^2 + x_{n+2}^2 + \cdots < \tfrac{1}{2}\varrho^2,$$

ferner wegen der Dichtigkeit der Menge der rationalen Punkte in E_n die rationalen Zahlen $r_1, r_2, \ldots, r_n$ so, daß

$$(x_1 - r_1)^2 + (x_2 - r_2)^2 + \cdots + (x_n - r_n)^2 < \tfrac{1}{2}\varrho^2.$$

Durch Addition folgt dann für $r = (r_1, \ldots, r_n, 0, 0, \ldots)$

$$f(x, r) < \varrho^2, \quad \overline{xr} < \varrho,$$

d. h. in jeder Umgebung U_x liegt ein Punkt von R. Als metrischer Raum mit einer abzählbaren dichten Teilmenge erfüllt E also auch die beiden Abzählbarkeitsaxiome (nach Ersetzung der sphärischen Umgebungen durch gleichwertige) und alles Bisherige bleibt in Geltung.

Will man die Beschränkung auf Folgen mit konvergenter Quadratsumme vermeiden, so wird man wieder für den Fall der Konvergenz von $f(x, y) = (x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots$ die Entfernung als $\overline{xy} = \sqrt{f(x, y)}$ definieren und gelangt dadurch zu einem teilweise metrischen Raum im Sinne von I. Alles bleibt richtig bis auf die Folgerungen aus dem zweiten Abzählbarkeitsaxiom, das hier nicht erfüllt ist.

In anderer Weise, nämlich vermöge der Definition der Entfernung durch die stets konvergente Reihe

$$\overline{xy} = \sum \frac{1}{n!} \frac{|x_n - y_n|}{1 + |x_n - y_n|},$$

hat M. Fréchet alle Folgen zu einem metrischen Raum vereinigt, in dem wieder die oben definierte abzählbare Menge R dicht ist. Eine ebenfalls zulässige Definition der Entfernung für $x \neq y$ ist der reziproke Wert $\frac{1}{n}$ der ersten Differenzstelle beider Zahlenfolgen; dies kann auf Elementfolgen aus einer beliebigen Menge ausgedehnt werden.

An Modifikationen und Verallgemeinerungen, die sich hier noch anschließen könnten, herrscht kein Mangel; wir schlagen dem Leser einige zur Übung vor:

Zahlenfolgen mit nur endlich vielen von Null verschiedenen Zahlen oder Punkte des Hilbertschen Raumes mit schließlich verschwindenden Koordinaten.

Zahlenkomplexe $x = (x_0, x_1, \ldots, x_\omega, \ldots)$ vom Typus ω_1, unter der Bedingung, daß höchstens abzählbar viele Koordinaten von Null verschieden sind und ihre Quadratsumme konvergiert; das Quadrat der

---

[1] Die Menge der Folgen rationaler Zahlen, ohne die letzte Bedingung, ist nicht abzählbar, sondern von der Mächtigkeit des Kontinuums ($\aleph_0^{\aleph_0} = \aleph$).

Entfernung sei wieder durch $(x_0 - y_0)^2 + (x_1 - y_1)^2 + \dots + (x_\omega - y_\omega)^2 + \dots$ definiert. In diesem „euklidischen Raum von $\aleph_1$ Dimensionen" ist keine abzählbare Menge, wohl aber die Menge R der Punkte mit nur endlich vielen nichtverschwindenden, rationalen Koordinaten dicht, die hier die Mächtigkeit $\aleph_1$ hat.

Zahlenfolgen mit kubischen Umgebungen: U_x sei die Menge der Folgen y, für die die obere Schranke der Beträge $|x_n - y_n|$ kleiner als eine positive Zahl ϱ ist.[1] Indem man diese Schranke, wenn sie existiert, als Entfernung $\overline{xy}$ definiert, erhält man einen teilweise metrischen Raum mit den sphärischen Umgebungen U_x. Offenbar läßt sich dies auf reelle Funktionen $x = x(t)$ einer beliebigen Variablen übertragen. Die Elemente $a_1, a_2, \dots$ konvergieren nach x, wenn die Funktionen $a_1(t), a_2(t), \dots$ gleichmäßig nach $x(t)$ konvergieren. (Vgl. das nächste Beispiel.)

Zahlenfolgen mit „Quadern" als Umgebungen: U_x sei durch die Ungleichungen $|x_n - y_n| < \varrho_n$ mit positiven $\varrho_1, \varrho_2, \dots$ definiert. Dies läßt sich auf Elementfolgen $x = (x_1, x_2, \dots)$ übertragen, deren Elemente x_n beliebige topologische Räume E_n durchlaufen; ist U_n eine Umgebung von x_n, so sei das Produkt $U_1 U_2 \dots$ (die Menge aller Folgen $y = (y_1, y_2, \dots)$, wo y_n Punkt von U_n ist) eine Umgebung von x. Auch hier kann man zu Funktionen einer beliebigen Variablen übergehen.

V. Stetige Funktionen.

Die Elemente x von E seien jetzt reelle Funktionen $x = x(t)$ einer reellen Variablen t und im Intervall $0 \leqq t \leqq 1$ stetig. Als Entfernung definieren wir

$$\overline{xy} = \max |x(t) - y(t)|,$$

das Maximum des Betrages der Differenz beider Funktionen im Intervall. Die Entfernungsaxiome sind erfüllt, also für die sphärischen Umgebungen U_x die Umgebungsaxiome. Es existiert wieder eine in E dichte abzählbare Menge R. Nach einem bekannten Satz von Weierstrass gibt es nämlich für jede solche Funktion $x(t)$ und jede positive Zahl ϱ ein Polynom

$$r(t) = a_0 + a_1 t + \dots + a_n t^n$$

derart, daß im ganzen Intervall $|x(t) - r(t)| < \varrho$; und wenn ein solches überhaupt existiert, so lehrt eine leichte Überlegung, daß man seine Koeffizienten rational annehmen darf. Die abzählbare Menge R der Polynome mit rationalen Koeffizienten ist also in E dicht; danach gelten für ein mit den sphärischen Umgebungen gleich-

[1] Wollte man diese Bedingung durch $|x_n - y_n| < \varrho$ ersetzen, so würden die Umgebungen das Axiom (C) nicht erfüllen (S. 213).

wertiges System die beiden Abzählbarkeitsaxiome und alle bisherigen Resultate bleiben richtig.

Eine andere zulässige Definition der Entfernung ist die nichtnegative Quadratwurzel aus

$$\int_0^1 [x(t) - y(t)]^2 dt;$$

bezeichnen wir sie mit $\widehat{xy}$ und die hiermit definierten sphärischen Umgebungen mit V_x (die U_x entsprechen den kubischen, die V_x den sphärischen Umgebungen des gewöhnlichen Raumes). Auf Grund einer bekannten Integralabschätzung ist für $\overline{xy} < \varrho$ auch $\widehat{xy} < \varrho$, also enthält jedes V_x ein U_x. Umgekehrt ist aber kein V_x in einem U_x enthalten, denn das Integral kann beliebig klein und doch der Integrand an einzelnen Stellen beliebig groß werden. Die U_x und V_x sind also nicht gleichwertig und definieren verschiedene Mengen A_α usw.; für beide aber bleiben alle bisherigen Sätze in Kraft, da auch bei Zugrundelegung der V_x die obengenannte Menge R in E dicht ist.

Ein weiteres Beispiel von der Art der in diesem Paragraphen behandelten lernen wir in § 6 kennen, indem wir die Teilmengen eines metrischen Raumes als Elemente eines neuen Raumes ansehen.

§ 6. Metrische Räume: Entfernungen und Zusammenhang.

Wir halten jetzt den Zeitpunkt für gekommen, wo eine weitere Fortsetzung der Umgebungstheorie mit einer Einbuße an Einfachheit verknüpft sein würde. Insbesondere hat ein metrischer Raum etwas, was sich rein topologisch nur umständlich beschreiben läßt, nämlich die durch gleichen Radius vermittelte Beziehung zwischen den Umgebungen verschiedener Punkte. In einem metrischen Raume, den wir von nun an voraussetzen, gelten die Umgebungsaxiome und (bei Ersetzung der Umgebungen durch gleichwertige) das erste Abzählbarkeitsaxiom, während das zweite die metrischen Räume mit abzählbarer dichter Teilmenge charakterisiert.

Eine Menge A heißt beschränkt, wenn die Entfernungen ihrer Punkte voneinander eine obere Schranke $d(A)$ haben; diese heißt die Breite von A. Endliche Mengen sind beschränkt; auch die Nullmenge rechnen wir zu den beschränkten Mengen, ohne ihr aber eine Breite zuzuschreiben. Die Teilmengen einer beschränkten Menge sind beschränkt; eine Summe endlich vieler beschränkter Mengen ist beschränkt. Kompakte Mengen sind beschränkt, denn eine unbeschränkte Menge hat offenbar divergente Teilmengen.

Die Entfernungen $\varrho = \overline{ab}$ der Punktpaare zweier nichtverschwindender Mengen A, B haben stets eine untere Schranke, die **untere Entfernung** zwischen A, B

$$\delta(A, B) = \delta(B, A) \geqq 0,$$

und, falls beide Mengen beschränkt sind, eine obere Schranke, die **obere Entfernung**

$$d(A, B) = d(B, A).$$

Die obere Entfernung einer Menge von sich selbst ist ihre Breite $d(A, A) = d(A)$. Die untere und obere Schranke der Entfernungen $\overline{ab}$, wenn a die Menge A durchläuft und b ein fester Punkt ist, bezeichnen wir mit

$$\delta(A, b) = \delta(b, A), \quad d(A, b) = d(b, A)$$

statt mit $\delta(A, \{b\})$ usw.[1]

Die oberen Abstände erfüllen die dem Dreiecksaxiom analoge Beziehung

$$(1) \qquad d(A, B) + d(B, C) \geqq d(A, C).$$

Denn wählt man a und c so, daß $\overline{ac} > d(A, C) - \varepsilon$, bei beliebig vorgeschriebenem positivem ε, so ist für jeden Punkt b

$$d(A, B) + d(B, C) \geqq \overline{ab} + \overline{bc} \geqq \overline{ac} > d(A, C) - \varepsilon,$$

$$d(A, B) + d(B, C) > d(A, C) - \varepsilon,$$

und aus der letzten, für jedes positive ε gültigen Ungleichung folgt (1).

Die unteren Abstände erfüllen die Dreiecksrelation nicht (z. B. wähle man A und C mit positiver unterer Entfernung, B als eine beide Mengen treffende Menge, so daß $\delta(A, B) = \delta(B, C) = 0$, $\delta(A, C) > 0$). Es gilt nur, wenn die mittlere Menge sich auf einen Punkt reduziert:

$$(2) \qquad \delta(A, b) + \delta(b, C) \geqq \delta(A, C).$$

Denn wählt man a und c so, daß $\overline{ab} < \delta(A, b) + \varepsilon$ und $\overline{bc} < \delta(b, C) + \varepsilon$, so ist

$$\delta(A, b) + \delta(b, C) + 2\varepsilon > \overline{ab} + \overline{bc} \geqq \overline{ac} \geqq \delta(A, C).$$

Spezialfälle von (1) und (2) sind

$$(3) \qquad d(A, b) + \overline{bc} \geqq d(A, c),$$

$$(4) \qquad \delta(A, b) + \overline{bc} \geqq \delta(A, c).$$

Die Formel (3) nebst der durch Vertauschung von b, c entstehenden besagt, daß der Betrag der Differenz $d(A, b) - d(A, c)$ höchstens

[1] Damit keine Kollision dadurch entstehe, ist wieder wie in früheren Fällen (S. 43, 85) vorauszusetzen, daß Elemente und Teilmengen von E, Punkte und Punktmengen verschiedene Dinge seien, abgesehen vielleicht von Gleichungen der Form $x = \{x\}$.

gleich $\overline{bc}$ ist. Wenn also b_n nach c konvergiert, so konvergiert $d(A, b_n)$ nach $d(A, c)$ und ebenso $\delta(A, b_n)$ nach $\delta(A, c)$; in der üblichen Ausdrucksweise (Kap. IX) sind $d(A, x)$ und $\delta(A, x)$ stetige Funktionen von x.

Es ist

$$(5) \qquad \delta(A, B) = \delta(A_\alpha, B), \quad d(A, B) = d(A_\alpha, B),$$

d. h. diese Entfernungen bleiben ungeändert, wenn man der Menge A ihre Häufungspunkte hinzufügt. Um dies z. B. für d zu beweisen: ist $\varepsilon > 0$ beliebig, x ein Punkt von A_α, a ein Punkt von A mit $\overline{ax} < \varepsilon$, endlich b ein beliebiger Punkt von B, so ist

$$\overline{xb} < \overline{ab} + \varepsilon \leqq d(A, B) + \varepsilon,$$

die $\overline{xb}$ sind also beschränkt (d. h. mit A ist A_α beschränkt) und $d(A_\alpha, B) \leqq d(A, B) + \varepsilon$, also $d(A_\alpha, B) \leqq d(A, B)$. Wegen $A_\alpha \geqq A$ ist andererseits $d(A_\alpha, B) \geqq d(A, B)$, also beide Zahlen gleich.

Man kann auch sagen: A und A' haben dieselben Entfernungen von B, wenn $A_\alpha = A'_\alpha$, wenn also (Kap. VII, § 8) die Mengen A, A' zu einander dicht oder, wie wir uns damals ausdrückten, kongruent sind ($A \equiv A'$). In den unteren und oberen Entfernungen kann jede Menge durch eine kongruente Menge ersetzt werden.

I. Sind die Mengen A, B abgeschlossen und kompakt, so bilden die Entfernungen $\overline{ab}$ ihrer Punktpaare eine abgeschlossene, beschränkte Zahlenmenge.

Denn ist P diese Menge und σ eine α-Zahl von P, so gibt es eine Folge von Punktpaaren mit $\lim \overline{a_n b_n} = \sigma$. Die Folge der a_n hat dann mindestens einen zu A gehörigen Häufungspunkt a, es ist also, für eine geeignete Folge natürlicher Zahlen p, $\lim a_p = a$, und ebenso für eine geeignete Teilfolge dieser Folge $\lim b_q = b$. Dann ist $\sigma = \lim \overline{a_q b_q} = \overline{ab}$, σ ist also eine wirklich erreichte Entfernung oder eine Zahl von P, womit diese Menge als abgeschlossen erkannt ist. Genau ebenso wird gezeigt, daß niemals $\lim \overline{a_n b_n} = +\infty$ sein kann, daß also P beschränkt ist, was wir übrigens schon wissen. Insbesondere existiert also in diesem Fall ein Minimum und ein Maximum der Entfernungen

$$\min \overline{ab} = \delta(A, B), \quad \max \overline{ab} = d(A, B),$$

die Entfernungsschranken werden wirklich erreicht.

Die Voraussetzung der Kompaktheit ist ebenso wesentlich wie die der Abgeschlossenheit. Ein Hyperbelzweig und eine seiner Asymptoten sind abgeschlossen; aber die untere Entfernung Null wird nicht erreicht. Es genügt auch nicht, daß beide Mengen

beschränkt und abgeschlossen seien (vgl. § 8). Im Hilbertschen Raume (S. 287) sei A die Menge der Punkte

$$a_1 = (1, 0, 0, 0, \ldots)$$
$$a_2 = (0, 1, 0, 0, \ldots)$$
$$a_3 = (0, 0, 1, 0, \ldots) \quad \text{usw.};$$

ist $b = (\beta_1, \beta_2, \beta_3, \ldots)$ ein weiterer Punkt mit lauter positiven Koordinaten und $\beta^2 = \beta_1{}^2 + \beta_2{}^2 + \beta_3{}^2 + \ldots$, so ist

$$\overline{a_n b}^2 = 1 - 2\beta_n + \beta^2$$

und diese Zahlen haben die obere Schranke $1 + \beta^2$, erreichen sie aber nicht. Sind die β_n sämtlich negativ, so erreichen die genannten Zahlen ihre untere Schranke $1 + \beta^2$ nicht. Die Menge A ist beschränkt und abgeschlossen, aber nicht kompakt (sie ist divergent).

Im Anschluß an das Bisherige erhebt sich die Frage, ob es nicht möglich ist, den (nichtverschwindenden) Teilmengen des metrischen Raumes E Entfernungen $\overline{AB}$ zuzuordnen, die auch ihrerseits die Entfernungsaxiome erfüllen. Hierzu eignen sich weder die unteren Entfernungen, die das Dreiecksaxiom verletzen, noch die oberen, wo $d(A, A)$ im allgemeinen nicht Null ist. Man gelangt aber zu solchen Entfernungen in folgender Weise.

Wir betrachten den unteren Abstand $\delta(A, b)$ und lassen b die Menge B durchlaufen. Wenn diese Abstände eine obere Schranke haben[1], so bezeichnen wir diese mit $\overrightarrow{AB}$; hierbei kommt es aber auf die Reihenfolge an, und die Zahlen $\overrightarrow{AB}$ und $\overrightarrow{BA}$ können, wenn sie beide existieren, verschieden sein. Um das Symmetrieaxiom zu sichern, definieren wir als Entfernung $\overline{AB}$ die größere[2] der beiden Zahlen

$$\overline{AB} = \max \overrightarrow{AB}, \overrightarrow{BA},$$

vorausgesetzt, daß beide existieren.

Die Entfernung liegt offenbar zwischen der unteren und oberen Entfernung; d. h. wenn $\overline{AB}$ existiert, so ist $\overline{AB} \geqq \delta(A, B)$, und wenn $d(A, B)$ existiert, so existiert auch $\overline{AB} \leqq d(A, B)$. Zwei beschränkte Mengen haben stets eine Entfernung, zwei unbeschränkte möglicherweise, eine beschränkte und eine unbeschränkte niemals.

Wenn $B = \{b\}$ aus einem einzigen Punkte besteht, so ist $\overrightarrow{Ab} = \delta(A, b)$, $\overrightarrow{bA} = d(A, b)$, also $\overline{Ab} = d(A, b)$; die Entfernung ist in diesem Falle die obere Entfernung.

---

[1] Dazu ist die Beschränktheit von A und B nicht notwendig, wie der Fall von zwei parallelen Geraden lehrt.

[2] Auch ihr arithmetisches Mittel würde im wesentlichen dieselben Dienste leisten.

Wenn $\overrightarrow{AB} < \varrho$, so gibt es zu jedem Punkte b einen Punkt a mit $\overline{ab} < \varrho$; wenn es umgekehrt zu jedem Punkte b einen Punkt a mit $\overline{ab} < \varrho$ gibt, so ist $\overrightarrow{AB} \leqq \varrho$. Diese leicht beweisbare Bemerkung, die auch zur Definition von $\overrightarrow{AB}$ dienen könnte, ist für Anwendungen besonders zweckmäßig.

Wenn $\overrightarrow{AB}$ und $\overrightarrow{BC}$ existieren, so existiert auch $\overrightarrow{AC}$ und es ist

$$\overrightarrow{AB} + \overrightarrow{BC} \geqq \overrightarrow{AC}. \tag{6}$$

Denn setzen wir $\overrightarrow{AB} = \varrho$, $\overrightarrow{BC} = \sigma$, so existiert für beliebiges positives ε zu jedem Punkt c ein Punkt b mit $\overline{bc} < \sigma + \varepsilon$, dazu ein Punkt a mit $\overline{ab} < \varrho + \varepsilon$, also $\overline{ac} < \varrho + \sigma + 2\varepsilon$, folglich ist

$$\overrightarrow{AC} \leqq \varrho + \sigma + 2\varepsilon \quad \text{und} \quad \overrightarrow{AC} \leqq \varrho + \sigma.$$

Hieraus folgt unmittelbar: wenn $\overline{AB}$ und $\overline{BC}$ existieren, so existiert auch $\overline{AC}$ und es ist

$$\overline{AB} + \overline{BC} \geqq \overline{AC}. \tag{7}$$

Denn die linke Seite ist

$$\geqq \overrightarrow{AB} + \overrightarrow{BC} \geqq \overrightarrow{AC} \quad \text{und} \quad \geqq \overrightarrow{CB} + \overrightarrow{BA} \geqq \overrightarrow{CA}.$$

Die Entfernungen erfüllen also das Dreiecksaxiom eines teilweise metrischen Raumes (S. 285).

Das Koinzidenzaxiom ist zunächst nicht erfüllt; zwar ist $\overline{AA} = 0$, aber $\overline{AB}$ kann auch noch sonst verschwinden. Die Gleichung $\overrightarrow{AB} = 0$ besagt, daß, für jeden Punkt b, $\delta(A, b) = 0$ ist (oder zu jedem b gibt es bei beliebigem positivem ϱ ein a mit $\overline{ab} < \varrho$), d. h. daß b stets α-Punkt von A, A zu B dicht ist, wovon offenbar auch die Umkehrung gilt. Also verschwindet $\overline{AB}$ dann und nur dann, wenn A und B zu einander dicht (kongruent) sind, woraus (nach dem Dreiecksaxiom) folgt, daß ohne Änderung der Entfernung jede Menge durch eine kongruente ersetzt werden kann. Um das Koinzidenzaxiom aufrechtzuerhalten, müssen wir also entweder kongruente Mengen als nicht verschieden ansehen (mit erweiterter Bedeutung des Begriffs Gleichheit) oder dürfen aus jeder Klasse kongruenter Mengen nur eine Menge, am einfachsten die abgeschlossene (größte) Menge der Klasse, zulassen. Wir können also sagen:

II. Die nichtverschwindenden abgeschlossenen Teilmengen eines metrischen Raumes E bilden ihrerseits einen teilweise metrischen Raum $\mathfrak{E}$.

Das System der beschränkten abgeschlossenen Mengen ist einer der metrischen Teilräume, in die der Mengenraum $\mathfrak{E}$ zerfällt. Hiervon bilden die kompakten abgeschlossenen Mengen wieder einen metrischen Teilraum $\mathfrak{C}$. Wenn in E eine abzählbare Menge R

dicht ist, so ist auch in $\mathfrak{C}$ eine abzählbare Menge dicht, nämlich das System der endlichen Teilmengen von R. In der Tat: ist A abgeschlossen und kompakt, ϱ eine positive Zahl, und schließen wir jeden Punkt a von A in seine Umgebung U_a mit dem Radius ϱ ein, so ist nach dem Borelschen Satze A bereits in einer endlichen Zahl dieser Umgebungen enthalten, deren Mittelpunkte $a_1, a_2, \ldots, a_n$ sein mögen. Zu jedem dieser Punkte a_i $(i = 1, 2, \ldots, n)$ gibt es einen Punkt r_i von R mit $\overline{a_i r_i} < \varrho$: die Punkte r_i bilden eine endliche Teilmenge R_0 von R. Zu jedem Punkt r_i von R_0 gibt es einen Punkt a_i von A mit $\overline{r_i a_i} < \varrho$, also ist $\overrightarrow{AR_0} < \varrho$. Zu jedem Punkt a von A gibt es einen Punkt a_i mit $\overline{a a_i} < \varrho$ und einen Punkt r_i mit $\overline{a r_i} < 2\varrho$, demnach ist $\overrightarrow{R_0 A} \leqq 2\varrho$. Also ist $\overline{AR_0} \leqq 2\varrho$, in beliebiger Nähe eines A liegt ein R_0.

Im ganzen Mengenraum $\mathfrak{E}$ ist im allgemeinen keine abzählbare Menge dicht, selbst wenn E diese Eigenschaft hat. Ist z. B. E die euklidische Ebene, so haben schon die verschiedenen Geraden durch einen Punkt paarweise keine Entfernung voneinander[1], und $\mathfrak{E}$ zerfällt mindestens in $\aleph$ verschiedene metrische Räume (übrigens auch in nicht mehr, da es überhaupt nur $\aleph$ abgeschlossene Mengen gibt, § 3, IV). Eine in $\mathfrak{E}$ dichte Menge hat daher die Mächtigkeit des Kontinuums.

Betrachten wir eine Folge von nichtverschwindenden Mengen $A_1, A_2, \ldots$ und erinnern uns aus Kap. VII, § 5 der beiden abgeschlossenen Limites

$$L = \overline{\text{Lim}} \inf A_n, \quad M = \overline{\text{Lim}} \sup A_n,$$

die wir im Falle der Gleichheit als abgeschlossenen Limes $L = M = \overline{\text{Lim}}\, A_n$ bezeichnet haben.[2] Andererseits sind auf Grund der Entfernungen zwischen Mengen Häufungselemente X und ein etwaiger Limes X der Mengenfolge wie gewöhnlich zu definieren, nämlich dadurch, daß, bei beliebigem $\varrho > 0$, die Ungleichung $\overline{XA_n} < \varrho$ für unendlich viele resp. fast alle n erfüllt ist; wir wollen dabei das durch jede kongruente Menge ersetzbare X durch die Forderung der Abgeschlossenheit präzisieren. . Den Zusammenhang dieser Limesbegriffe vermittelt der Satz:

III. Wenn $\lim \overrightarrow{A_n X} = 0$, so ist $X \leqq L$; wenn $\lim \overrightarrow{YA_n} = 0$, so ist $Y \geqq M$ (X, Y als abgeschlossen vorausgesetzt).

[1] Bei dieser Entfernungsdefinition konvergiert also im allgemeinen die Folge der Geraden $y = a_n x + b_n$ nicht gegen die Gerade $y = ax + b$, wenn a_n nach a und b_n nach b konvergiert; nur parallele Gerade haben eine Entfernung.

[2] Wir lassen die Striche über L, M jetzt weg, da von den übrigen Limites hier nicht die Rede ist.

Denn für beliebiges positives ϱ ist, für fast alle n, $\overrightarrow{A_n X} < \varrho$, so daß es zu jedem Punkt x von X einen Punkt a_n von A_n mit $\overline{xa_n} < \varrho$ gibt, d. h. jede Umgebung von x hat Punkte mit fast allen A_n gemein, x ist Punkt von L, $X \leqq L$. Andererseits sei m ein Punkt von M; jede Umgebung von m hat Punkte mit unendlich vielen A_n gemein oder es gibt, für unendlich viele n, einen Punkt a_n von A_n mit $\overline{ma_n} < \varrho$. Wählt man unter diesen ein n so groß, daß $\overrightarrow{YA_n} < \varrho$, so gibt es zu a_n einen Punkt y mit $\overline{ya_n} < \varrho$, also $\overline{my} < 2\varrho$. Demnach ist m ein α-Punkt von Y; $M \leqq Y_\alpha = Y$.

Ist jetzt $X = \lim A_n$, so ist $\lim \overline{XA_n} = 0$, also nach III

$$X \leqq L \leqq M \leqq X, \quad L = M = X:$$

IV. **Wenn die Mengenfolge nach X konvergiert, so ist X auch ihr abgeschlossener Limes.**

Umgekehrt läßt sich nicht soviel aussagen. Ist X ein Häufungselement der Mengenfolge, so gibt es nach dem hier gültigen Satz von der Trennbarkeit der Häufungspunkte (§ 2) eine nach X konvergente Teilfolge A_p, deren abgeschlossene Limites also mit X identisch sind. Daher ist (S. 237) $L \leqq X \leqq M$, alle Häufungselemente der Folge liegen zwischen L und M. **Wenn insbesondere die Mengenfolge einen abgeschlossenen Limes $L = M$ hat, so kann sie kein von diesem verschiedenes Häufungselement haben.** Aber daraus läßt sich ihre Konvergenz natürlich im allgemeinen nicht erschließen.

Es gibt aber einen Fall, wo man mehr aussagen kann, wenn nämlich die Mengen A_n Teilmengen einer und derselben **kompakten** Menge sind oder, was dasselbe ist, wenn ihre Summe $S = \mathfrak{S}(A_1, A_2, \ldots)$ kompakt ist. Es gilt nämlich folgendes Gegenstück zu III:

V. **Wenn die Mengen A_n eine kompakte Summe haben, so ist $\lim \overrightarrow{MA_n} = 0$ und, wenn L nicht verschwindet, $\lim \overrightarrow{A_n L} = 0$.**

Mit S ist (nach S. 234) auch S_α und deren Teilmengen L, M, A_n kompakt. M ist sicher von Null verschieden, da eine beliebige Folge von Punkten a_n aus A_n sicher einen Häufungspunkt hat, der nach Definition zu M gehört (die A_n waren ja > 0 angenommen). Alle Entfernungen $\overline{MA_n}$ und, wenn $L > 0$, alle Entfernungen $\overline{LA_n}$ sind vorhanden.

Wäre nun nicht $\lim \overrightarrow{MA_n} = 0$, so wäre für ein gewisses positives ϱ, und für unendlich viele n, $\overrightarrow{MA_n} > \varrho$, also für geeignete Punkte a_n $\delta(M, a_n) > \varrho$. Diese Punkte a_n haben aber einen Häufungspunkt m in M, für unendlich viele von ihnen ist also $\overline{ma_n} < \varrho$ und $\delta(M, a_n) < \varrho$, was ein Widerspruch ist.

Wäre zweitens, für $L > 0$, nicht $\lim \overrightarrow{A_n L} = 0$, so wäre für ein gewisses ϱ und, für unendlich viele n, $\overrightarrow{A_n L} > 2\varrho$, also für geeignete Punkte l_n von L, $\delta(A_n, l_n) > 2\varrho$. Diese Punkte l_n haben aber einen Häufungspunkt l in L, und für unendlich viele von ihnen ist $\overline{l l_n} < \varrho$, also $\delta(A_n, l) \geqq \delta(A_n, l_n) - \overline{l l_n} > \varrho$. Dann hätte aber die Umgebung von l mit dem Radius ϱ mit unendlich vielen A_n keinen Punkt gemein, im Widerspruch zur Definition von L.

Aus V folgt nun:

VI. **Wenn die Mengenfolge eine kompakte Summe und einen abgeschlossenen Limes $L = M$ hat, so konvergiert sie nach diesem.**

Beispiele (vgl. Kap. VII, § 5).

Die Folge $A, B, A, B, \ldots$, wo wir A und B abgeschlossen und > 0 voraussetzen, gibt $L = \mathfrak{D}(A, B)$, $M = \mathfrak{S}(A, B)$. Zur Konvergenz ist $L = M$, $A = B$ notwendig, aber auch hinreichend.

Im Beispiel (η) a. a. O. ist die Mengenfolge divergent; denn ihr einziges Häufungselement könnte nur das Paar der Halbgeraden sein, aber dieses hat von keinem A_n eine Entfernung. Man sieht hieran, daß der Satz VI nicht richtig bleibt, wenn nur die einzelnen A_n kompakt sind.

A_n sei die Ordinatenachse, gekreuzt von der horizontalen Strecke $y = n$, $-1 \leqq x \leqq 1$. Hier ist $L = M$ die Ordinatenachse; sie hat von jedem A_n die Entfernung 1; die Mengenfolge ist divergent.



Fig. 10.

Parallele Gerade in den Abständen $1, \frac{1}{2}, \frac{1}{3}, \ldots$ von einer festen Geraden konvergieren gegen diese Gerade; die einem Kreis eingeschriebenen oder umschriebenen regulären Polygone konvergieren gegen den Kreis usw.

Um noch eine Folgerung aus V zu ziehen, bemerken wir zunächst die Ungleichung

$$d(B, B') - d(A, A') \leqq \overrightarrow{AB} + \overrightarrow{A'B'},$$

die man so beweist. Man wähle, für vorgegebenes positives ε, zwei Punkte b, b' mit $\overline{bb'} > d(B, B') - \varepsilon$, sodann zwei Punkte a, a' mit

$$\overline{ab} < \overrightarrow{AB} + \varepsilon, \quad \overline{a'b'} < \overrightarrow{A'B'} + \varepsilon,$$

dann ist

$$\overrightarrow{AB} + \overrightarrow{A'B'} + d(A, A') + 3\varepsilon$$

$$> \overline{ab} + \overline{a'b'} + \overline{aa'} + \varepsilon \geqq \overline{bb'} + \varepsilon > d(B, B').$$

Genau entsprechend folgt

$$\delta(A, A') - \delta(B, B') \leqq \overrightarrow{AB} + \overrightarrow{A'B'},$$

wenn man zuerst $\overline{bb'} < \delta(B, B') + \varepsilon$ wählt und weiter wie soeben verfährt.

Daraus ergibt sich:

VII. Wenn die Mengen A_n und A_n' kompakte Summen haben, so ist

$$d(M, M') \geqq \limsup d(A_n, A_n'), \quad \delta(M, M') \leqq \liminf \delta(A_n, A_n')$$

und, falls L, L' nicht verschwinden,

$$d(L, L') \leqq \liminf d(A_n, A_n'), \quad \delta(L, L') \geqq \limsup \delta(A_n, A_n').$$

Man hat nämlich

$$d(A_n, A_n') - d(M, M') \leqq \overrightarrow{MA_n} + \overrightarrow{M'A_n'}$$

und daraus (durch den Schluß von $x_n \leqq y_n$ auf $\limsup x_n \leqq \limsup y_n$), weil die rechte Seite nach V den lim = lim sup Null hat, die erste der angegebenen Formeln; genau so folgen die übrigen.

Wenn die abgeschlossenen Limites $L = M = X$, $L' = M' = X'$ existieren, so folgt daraus

$$d(X, X') = \lim d(A_n, A_n'), \quad \delta(X, X') = \lim \delta(A_n, A_n'),$$

z. B. für konstantes, abgeschlossenes $A_n' = X'$

$$d(X, X') = \lim d(A_n, X'), \quad \delta(X, X') = \lim \delta(A_n, X'),$$

oder, wenn man beide Mengenfolgen identifiziert ($A_n' = A_n$),

$$d(X) = \lim d(A_n),$$

die Breite von A_n konvergiert nach der Breite des abgeschlossenen Limes.

Die Theorie des Zusammenhanges (Kap. VII, § 7) gestattet in einem metrischen Raume noch eine weitere Ausführung, nämlich eine Art gradueller Abstufung, bei der auch eine unzusammenhängende Menge noch einen gröberen Zusammenhang haben kann und erst bei schärferer Betrachtung in Teile zerfällt, wie sich ein Nebelfleck in Sterne auflöst. Wir wollen, für eine positive Zahl ϱ, eine nichtverschwindende Menge A ϱ-zusammenhängend nennen, wenn bei jeder Zerlegung $A = P + Q$ in zwei nichtverschwindende Summanden deren untere Entfernung $\delta(P, Q) \leqq \varrho$ ist, ebenso 0-zusammenhängend, wenn bei jeder solchen Zerlegung $\delta(P, Q) = 0$ ist (also A ϱ-zusammenhängend für jedes positive ϱ).[1] Eine schlechthin zusammenhängende Menge ist 0-zusammenhängend, da bei jeder Zerlegung $A = P + Q$ wenigstens die eine Menge einen Häufungspunkt der andern enthält; daß das Umgekehrte im allgemeinen nicht gilt, werden wir an Beispielen sehen.

VIII. Eine Summe von beliebig vielen ϱ-zusammenhängenden Mengen, die paarweise untere Entfernungen $\leqq \varrho$ haben, ist wieder ϱ-zusammenhängend. Eine Summe von

[1] Eine Menge aus einem Punkt ist als ϱ-zusammenhängend und 0-zusammenhängend zu betrachten.

beliebig vielen 0-zusammenhängenden Mengen, die paarweise die untere Entfernung 0 haben, ist 0-zusammenhängend.

Denn wird $A = \mathfrak{S} A_i$ in zwei nichtverschwindende Teilmengen $A = P + Q$ zerlegt, so wird entweder mindestens ein Summand $A_i = \mathfrak{D}(A_i, P) + \mathfrak{D}(A_i, Q) = P_i + Q_i$ ebenfalls in zwei nichtverschwindende Teilmengen zerlegt, und dann ist $\delta(P, Q) \leqq \delta(P_i, Q_i) \leqq \varrho$; oder jeder Summand bleibt ganz in einer der Mengen P, Q, und wenn dann $A_i \leqq P$, $A_k \leqq Q$, so ist $\delta(P, Q) \leqq \delta(A_i, A_k) \leqq \varrho$. Der zweite Satz ist unmittelbare Folge des ersten (oder wird genau so mit $\varrho = 0$ bewiesen).

Unter einer ϱ-**Komponente** resp. 0-**Komponente** von A verstehen wir eine größte (in keiner andern enthaltene) ϱ-zusammenhängende resp. 0-zusammenhängende Teilmenge von A. Die Existenz solcher geht aus VIII hervor: die Summe aller ϱ-zusammenhängenden Teilmengen von A, die den festen Punkt p enthalten, ist eine ϱ-Komponente. Zwei verschiedene ϱ-Komponenten $P \neq Q$ sind nicht nur fremd, sondern haben sogar eine untere Entfernung $\delta(P, Q) > \varrho$, da andernfalls $\mathfrak{S}(P, Q) > P$ oder $> Q$ und auch noch ϱ-zusammenhängend wäre; eine (nicht mit A selbst identische) ϱ-Komponente P hat von ihrem Komplement $A - P$ eine untere Entfernung $\geqq \varrho$, beide sind in A abgeschlossen. Zwei verschiedene 0-Komponenten haben positive untere Entfernung; eine 0-Komponente (nicht aber im allgemeinen ihr Komplement) ist in A abgeschlossen. Wir bezeichnen die den Punkt p enthaltende ϱ-Komponente und 0-Komponente mit P_ϱ und P_0, die gewöhnliche Komponente mit P, die Quasikomponente (S. 248) mit P_q; außerdem sei

$$P_{\varrho+0} = \underset{\sigma}{\mathfrak{D}} P_\sigma \quad (\sigma > \varrho), \qquad P_{0+0} = \underset{\sigma}{\mathfrak{D}} P_\sigma \quad (\sigma > 0)$$

die Menge der Punkte, die mit p gleichzeitig derselben σ-Komponente für jedes $\sigma > \varrho$ resp. für $\sigma > 0$ angehören. Da von den Eigenschaften: zusammenhängend, 0-, ϱ-, σ-zusammenhängend, für $0 < \varrho < \sigma$ jede die folgende nach sich zieht, so ist

$$P \leqq P_0 \leqq P_{0+0} \leqq P_\varrho \leqq P_{\varrho+0} \leqq P_\sigma;$$

für die obigen Durchschnitte kann man offenbar solche aus Mengenfolgen setzen, z. B. $P_{0+0} = \mathfrak{D}(P_1, P_{\frac{1}{2}}, P_{\frac{1}{3}}, \ldots)$.

Die Quasikomponente P_q war der Durchschnitt aller den Punkt p enthaltenden Mengen, die, zugleich mit ihrem Komplement, in A abgeschlossen sind; da zu ihnen die ϱ-Komponenten gehören, so ist $P \leqq P_q \leqq P_{0+0}$, während die Beziehung von P_q zu P_0 im allgemeinen fraglich bleibt.

Als eine (nicht unbedingt notwendige) Ergänzung dieser Be-

trachtungen stellen wir noch die Definition auf: A heiße $(\varrho-0)$-zusammenhängend[1], wenn bei jeder Zerlegung $A = P + Q$ die untere Entfernung $\delta(P, Q) < \varrho$ ausfällt (statt $\leqq \varrho$ wie beim ϱ-Zusammenhang). Wie oben folgt, daß die Summe beliebig vieler solcher Mengen, die paarweise untere Entfernungen $< \varrho$ haben, wieder $(\varrho-0)$-zusammenhängend ist, und daraus die Existenz von $(\varrho-0)$-Komponenten einer Menge A. Für die den Punkt p enthaltende $(\varrho-0)$-Komponente gilt

$$P_{\varrho-0} \leqq P_\varrho, \; P_{\varrho-0} = \mathfrak{S}_\pi P_\pi \qquad (0 < \pi < \varrho),$$

wovon nur die zweite Formel eines Beweises bedarf. Nennen wir die rechte Seite S, so ist jedenfalls $P_{\varrho-0} \geqq S$. Wäre aber $P_{\varrho-0} > S$, also $P_{\varrho-0} = S + T$, $\delta(S, T) < \varrho$, also für geeignete Punkte $\overline{st} < \varrho$, so gehört s einer gewissen Menge P_π und allen folgenden an; wir können daher annehmen, daß $\overline{st} \leqq \pi < \varrho$. Dann wäre aber $P_\pi + \{t\}$ noch π-zusammenhängend, P_π nicht die größte den Punkt p enthaltende π-zusammenhängende Teilmenge von A.

Beispiele. Ein Hyperbelzweig und eine Asymptote sind zusammenhängend und haben die untere Entfernung 0; ihre Summe ist 0-zusammenhängend, aber nicht zusammenhängend.

Ein Punktpaar mit der Entfernung 1 ist 1-zusammenhängend, aber nicht $(1-0)$-zusammenhängend. Die lineare Menge der Punkte mit den Abszissen

$$0, \tfrac{1}{2}, \tfrac{1}{2} + \tfrac{2}{3}, \tfrac{1}{2} + \tfrac{2}{3} + \tfrac{3}{4}, \ldots$$

ist $(1-0)$-zusammenhängend. Die Menge der Punkte $0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, 2$ ist 1-zusammenhängend, ohne daß sich der Punkt 2 mit einem andern durch eine 1-Kette verbinden ließe, was erst eintrifft, wenn man der Menge noch ihren Häufungspunkt 1 hinzufügt.

Die Menge der rationalen Zahlen ist 0-zusammenhängend, aber nicht zusammenhängend; ihre Komponenten und Quasikomponenten bestehen aus je einem einzelnen Element, es ist also $P_0 > P_q$.

Ist A die Summe der Rechtecke und der beiden Geraden S. 249, p ein

---

[1] Nennt man einen endlichen Punktkomplex $(a_0, a_1, a_2, \ldots, a_n)$ eine ϱ-Kette, wenn die Entfernungen konsekutiver Punkte $\overline{a_0 a_1}, \overline{a_1 a_2}, \ldots, \overline{a_{n-1} a_n}$ sämtlich $\leqq \varrho$, eine $(\varrho-0)$-Kette, wenn sie sämtlich $< \varrho$ sind, so wird sich der Leser leicht von der Richtigkeit folgender Sätze überzeugen: eine Menge A, in der je zwei Punkte a, b durch eine ϱ-Kette $(a, \ldots, b)$ von Punkten in A verbunden werden können, ist ϱ-zusammenhängend (aber im allgemeinen nicht umgekehrt); eine Menge, in der je zwei Punkte durch eine $(\varrho-0)$-Kette verbunden werden können, ist $(\varrho-0)$-zusammenhängend, und umgekehrt. G. Cantor hat gelegentlich eine Menge zusammenhängend genannt, wenn je zwei ihrer Punkte, für jedes $\varrho > 0$, durch eine ϱ-Kette verbunden werden können; nach unserer Terminologie ist sie dann nur 0-zusammenhängend.

Punkt der einen Geraden, so ist $P_{0+0} = P_q$ das Geradenpaar, $P_0 = P$ die durch p gehende Gerade. Hier ist $P_0 \subset P_q$; außerdem $P_0 \subset P_{0+0}$, d. h. zwei Punkte, die für jedes ϱ derselben ϱ-Komponente angehören, brauchen nicht derselben 0-Komponente anzugehören.

Eine Vereinfachung erfahren diese Verhältnisse, wenn es sich um eine kompakte abgeschlossene Menge A handelt. Eine solche ist, wenn 0-zusammenhängend, auch zusammenhängend; denn ist sie unzusammenhängend und in zwei abgeschlossene Summanden > 0 zerlegbar, so ist deren untere (in diesem Fall wirklich erreichte) Entfernung positiv, und für jedes kleinere ϱ die Menge schon nicht mehr ϱ-zusammenhängend. Die Anzahl der ϱ-Komponenten ist für jedes ϱ endlich (übertrifft ϱ die Breite von A, so ist A ϱ-zusammenhängend); denn wählt man aus jeder einen beliebigen Punkt aus, so haben diese Punkte paarweise Entfernungen $> \varrho$, also keinen Häufungspunkt, und dürfen danach nur in endlicher Menge vorhanden sein. Es gibt nur endlich viele Zerlegungen $A = P + Q$ mit $\delta(P, Q) > \varrho$, denn bei jeder solchen bleiben die ϱ-Komponenten unzerlegt, und die Zerlegung ist nur eine Einteilung der ϱ-Komponenten in zwei Klassen. Bei allen Zerlegungen einer unzusammenhängenden Menge haben also die Zahlen $\delta(P, Q)$ ein Maximum $\varrho > 0$, und die Menge ist dann noch ϱ-zusammenhängend, aber nicht mehr $(\varrho - 0)$-zusammenhängend; sie ist σ-zusammenhängend für $\sigma \geqq \varrho$ und nicht π-zusammenhängend für $\pi < \varrho$. Eine $(\varrho - 0)$-zusammenhängende Menge ist auch noch für gewisse Zahlen $\pi < \varrho$ π-zusammenhängend, so daß für die Komponenten $P_\pi = P_{\varrho - 0}$ gilt.

Vor allem aber läßt sich jetzt unter gewissen Bedingungen aus dem Zusammenhang von Mengen einer Folge auf den Zusammenhang ihrer oberen abgeschlossenen Limes schließen. Es gilt der Satz:

IX. Wenn die ϱ-zusammenhängenden Mengen A_n eine kompakte Summe und einen nichtverschwindenden unteren abgeschlossenen Limes L haben, so ist ihr oberer abgeschlossener Limes M gleichfalls ϱ-zusammenhängend.

Nehmen wir an, es sei M nicht ϱ-zusammenhängend, also $M = P + Q$ mit $\delta(P, Q) = \sigma > \varrho$; wir setzen $\sigma - \varrho = 3\varepsilon$. Ein Punkt von L, der etwa zu P gehören möge, sei mit p bezeichnet, q ein beliebiger Punkt von Q. Die Umgebung U_p mit dem Radius ε hat mit fast allen, die Umgebung U_q mit dem Radius ε mit unendlich vielen A_n Punkte gemein; außerdem ist nach V, für fast alle n, $\overrightarrow{MA_n} < \varepsilon$; für unendlich viele A_n, deren eines mit A bezeichnet sei, treffen also alle drei Aussagen zu. A hat also zwei Punkte a, b mit $\overline{ap} < \varepsilon$, $\overline{bq} < \varepsilon$, und es ist $\overrightarrow{MA} < \varepsilon$.

Nun ist der ϱ-Zusammenhang von A zu berücksichtigen. Wir teilen die Punkte von A in drei Klassen:

(α) die Menge X der Punkte x, für die $\delta(P,x) < \varepsilon$, also nach (2)

$$\delta(Q,x) \geqq \delta(P,Q) - \delta(P,x) > \sigma - \varepsilon = \varrho + 2\varepsilon;$$

hierzu gehört der Punkt a.

(β) die Menge Y der Punkte y, für die $\delta(Q,y) < \varepsilon$, also in gleicher Weise $\delta(P,y) > \varrho + 2\varepsilon$; hierzu gehört b.

(γ) die Menge Z der übrigen Punkte z, für die also gleichzeitig $\delta(P,z) \geqq \varepsilon$ und $\delta(Q,z) \geqq \varepsilon$, also $\delta(M,z) \geqq \varepsilon$.

Auch solche Punkte z muß es geben. Denn je zwei Punkte x, y haben einen Abstand $\overline{xy} \geqq \delta(P,y) - \delta(P,x) > \varrho + \varepsilon$, also ist $\delta(X,Y) \geqq \varrho + \varepsilon > \varrho$, und wenn $Z = 0$ wäre, so wäre $A = X + Y$ nicht ϱ-zusammenhängend. Wenn es aber solche Punkte gibt, so ist $\overrightarrow{MA} \geqq \delta(M,z) \geqq \varepsilon$, im Widerspruch zur obigen Feststellung $\overrightarrow{MA} < \varepsilon$. Damit ist der Satz bewiesen.

X. Wenn die zusammenhängenden Mengen A_n eine kompakte Summe und einen nichtverschwindenden unteren abgeschlossenen Limes L haben, so ist ihr oberer abgeschlossener Limes M gleichfalls zusammenhängend.

Denn nach IX ist M, für jedes ϱ, ϱ-zusammenhängend, also 0-zusammenhängend und, als abgeschlossene kompakte Menge, zusammenhängend. (Man kann auch in dem obigen Beweise direkt $\varrho = 0$ annehmen.)

Daß dieser Satz ohne die Kompaktheit der Mengensumme, selbst wenn die einzelnen A_n kompakt sind, nicht zuzutreffen braucht, zeigen die Streckenzüge S. 239 oder die Rechtecke S. 249. Daß er für $L = 0$ nicht zu gelten braucht, zeigt schon die Folge $A, B, A, B, \ldots$, wo (bei abgeschlossenen A, B) $L = \mathfrak{D}(A,B)$ und $M = \mathfrak{S}(A,B)$; dies Beispiel zeigt auch, daß der Zusammenhang von L nicht behauptet werden kann.

Wenn eine Folge zusammenhängender Mengen mit kompakter Summe einen abgeschlossenen Limes hat, oder wenn der Limes $X = \lim A_n$ existiert (wobei, wie erinnerlich, X als abgeschlossen vorausgesetzt wurde), so ist dieser zusammenhängend. Jedes Häufungselement einer Folge zusammenhängender Mengen mit kompakter Summe ist zusammenhängend.

Diese Sätze erweitern den Kreis zusammenhängender Mengen durch Limesbildung; sie sind auch die Grundlage für diejenige Behandlungsweise der euklidischen Punktmengen, die sich auf approximierende Streckenzüge u. dgl. stützt.

Für eine absteigende Folge $A_1 \geqq A_2 \geqq \ldots$ von abgeschlossenen, kompakten, nichtverschwindenden Mengen ist $L = M = A = \mathfrak{D}(A_1, A_2, \ldots)$; sind die A_n ϱ-zusammenhängend, so ist ihr Durchschnitt ϱ-zusammen-

hängend. Sind die A_n ϱ_n-zusammenhängend und $\lim \varrho_n = \varrho$, so ist der Durchschnitt ϱ-zusammenhängend; denn für jedes $\sigma > \varrho$ ist schließlich (für fast alle n) $\varrho_n < \sigma$ und $A = \mathfrak{D}(A_n, A_{n+1}, \ldots)$ σ-zusammenhängend. Ebenso folgt für $\lim \varrho_n = 0$, daß A zusammenhängend ist. Für die Komponenten einer abgeschlossenen kompakten Menge ergibt sich daraus, daß der früher (S. 299) betrachtete Durchschnitt $P_{\varrho+0}$ ϱ-zusammenhängend, also $P_{\varrho+0} = P_\varrho$, und daß ebenso $P_{0+0} = P_0 = P$ ist, d. h.

$$P = P_q = P_0 = P_{0+0}.$$

Bei einer kompakten abgeschlossenen Menge sind also Komponenten, Quasikomponenten und 0-Komponenten identisch, und die Punkte, die für jedes ϱ derselben ϱ-Komponente angehören, gehören derselben Komponente an.

Zwischen den Komponenten einer abgeschlossenen kompakten Menge A läßt sich nun noch in anderer Weise, als dies oben (S. 293) allgemein geschehen ist, eine Entfernung definieren. Zwei verschiedene Komponenten P, Q gehören nicht für jedes ϱ derselben ϱ-Komponente an, wohl aber für hinlänglich großes ϱ; es gibt also eine trennende Zahl ϱ, derart, daß

$$P_\sigma = Q_\sigma \text{ für } \sigma > \varrho, \quad P_\pi \neq Q_\pi \text{ für } \pi < \varrho.$$

Dann ist auch noch $P_{\varrho+0} = Q_{\varrho+0}$ oder $P_\varrho = Q_\varrho$, und ϱ ist die kleinste Zahl, für welche P, Q noch derselben ϱ-Komponente angehören. Diese Zahl bezeichnen wir mit

$$\varrho = \underline{PQ} > 0,$$

während wir $\underline{PP} = 0$ setzen, und nennen sie etwa die Distanz[1] zwischen P und Q. Es ist leicht zu sehen, daß diese Distanzen die Entfernungsaxiome erfüllen; von dem Koinzidenz- und Symmetrieaxiom ist es ohne weiteres klar, und wenn P, Q, R drei Komponenten von A sind, die wir als sämtlich verschieden ansehen können, ferner ϱ das Maximum der beiden Distanzen $\underline{PQ}$, $\underline{QR}$ ist, so ist $P_\varrho = Q_\varrho = R_\varrho$, also $\underline{PR} \leqq \varrho$ und um so mehr

$$\underline{PQ} + \underline{QR} \geqq \underline{PR},$$

also das Dreiecksaxiom erfüllt.

Die Distanz ist höchstens gleich der unteren Entfernung: denn wenn $\delta(P, Q) = \varrho$, so ist $P + Q$ ϱ-zusammenhängend, P und Q gehören derselben ϱ-Komponente an, also $\underline{PQ} \leqq \varrho$ oder $\underline{PQ} \leqq \delta(P, Q)$.

---

[1] Die Distanz hängt auch von der Menge A ab: wenn P, Q zugleich Komponenten einer andern abgeschlossenen kompakten Menge A' sind, so können sie in dieser eine andere Distanz haben. In $P + Q$ haben sie die Distanz $\delta(P, Q)$.

Für eine Folge von Komponenten P_n wähle man nun aus jeder einen Punkt p_n; die Folge dieser Punkte hat in der kompakten abgeschlossenen Menge A gewiß einen Häufungspunkt p, der der Komponente P angehören möge; für eine passende Teilfolge natürlicher Zahlen m ist dann $\lim \overline{pp_m} = 0$, $\lim \delta(P, P_m) = 0$, $\lim \overline{PP_m} = 0$. Das heißt aber, daß es zu jeder Komponentenfolge eine Komponente gibt, die (im Sinne der Distanzen) Häufungselement der Folge ist, also:

XI. Die Menge der Komponenten einer abgeschlossenen kompakten Menge ist selbst, im Sinne der Distanzen, eine abgeschlossene kompakte Menge.[1]

Übrigens ist ein Häufungselement im Sinne der Distanzen auch eins im Sinne der unteren Entfernungen, d. h. mit $\overline{PP_n}$ konvergiert auch $\delta(P, P_n)$ nach Null, wovon das Umgekehrte wegen $\overline{PQ} \leqq \delta(P, Q)$ trivial ist. Denn zu der Folge der Komponenten P_n gibt es, wie wir soeben sahen, eine Komponente Q mit $\lim \delta(Q, P_m) = \lim \overline{QP_m} = 0$, woraus wegen $\lim \overline{PP_m} = 0$ zunächst $\overline{PQ} = 0$, $Q = P$ folgt. Also ist $\lim \delta(P, P_m) = 0$, P ist im Sinne der unteren Entfernungen Häufungselement der Folge, und da dies auch für jede Teilfolge gilt, so muß $\lim \delta(P, P_n) = 0$ sein. Ist $\mathfrak{M} = \{P, Q, \ldots\}$ irgend eine Menge von Komponenten von A (nicht notwendig aller) mit der Summe $M = P + Q + \ldots$, so ist P also isoliertes resp. Häufungselement von $\mathfrak{M}$, je nachdem die Distanzen $\overline{PQ}$ oder die unteren Entfernungen $\delta(P, Q)$ von den übrigen Komponenten eine positive resp. verschwindende untere Schranke haben, d. h. je nachdem das Komplement $M - P$ in M abgeschlossen ist oder nicht. Diese Unterscheidung ist also, wenn $P, Q, \ldots$ gleichzeitig die Komponenten einer andern abgeschlossenen kompakten Menge A' sind, unabhängig davon, ob die Distanzen auf A oder A' bezogen werden.

§ 7. Metrische Räume: Borelsche Mengen.

Eine wichtige Besonderheit des metrischen Raumes ist die Darstellung abgeschlossener Mengen durch Folgen von Gebieten und umgekehrt. Wir erinnern hier zuvor an die formalen Betrachtungen von Kap. I, § 10; wie damals bezeichnen wir mit

$$A_\sigma = \mathfrak{S}(A_1, A_2, \ldots), \quad A_\delta = \mathfrak{D}(A_1, A_2, \ldots)$$

die Summe und den Durchschnitt einer Folge von Mengen A, die

[1] Im Sinne der Entfernungen $\overline{PQ}$ trifft dies nicht zu; die Häufungselemente von Komponentenfolgen können zusammenhängende Teilmengen von Komponenten sein, wie die einfachsten Beispiele lehren.

einem gewissen Mengensystem $\mathfrak{A}$ angehören; die A_σ bilden das kleinste σ-System $\mathfrak{A}_\sigma$, die A_δ das kleinste δ-System $\mathfrak{A}_\delta$ über $\mathfrak{A}$. Auch die Fortsetzung dieser Prozesse haben wir besprochen; z. B. bilden die Durchschnitte $A_{\sigma\delta}$ aus Folgen von Mengen A_σ das kleinste δ-System über $\mathfrak{A}_\sigma$. Die inzwischen erworbene Kenntnis der Ordnungszahlen gestattet uns auch, das kleinste System $\mathfrak{A}_{(\sigma\delta)}$ über $\mathfrak{A}$ zu bilden, das sowohl die Summe als auch den Durchschnitt jeder Folge seiner Mengen enthält: setzt man etwa $\varphi(\mathfrak{A}) = \mathfrak{A}_{\sigma\delta}$, und definiert für jede Ordnungszahl $\eta < \omega_1$ durch Induktion

$$\mathfrak{A}_0 = \mathfrak{A},\ \mathfrak{A}_\eta = \underset{\xi}{\mathfrak{S}}\varphi(\mathfrak{A}_\xi) \qquad (\xi < \eta),$$

so ist $\mathfrak{A}_{(\sigma\delta)} = \underset{\eta}{\mathfrak{S}}\mathfrak{A}_\eta$ $(\eta < \omega_1)$, was indessen nicht ausschließt, daß der Prozeß schon früher zum Abschluß kommt. Ist das System $\mathfrak{A}$ von der Mächtigkeit $\aleph$ des Kontinuums, was z. B. für die abgeschlossenen Mengen oder Gebiete eines Raumes mit zweitem Abzählbarkeitsaxiom zutrifft (§ 3, IV), so ist $\mathfrak{A}_\sigma$ von derselben Mächtigkeit $\aleph^{\aleph_0} = \aleph$, ebenso $\mathfrak{A}_{\sigma\delta}$ und schließlich das ganze System $\mathfrak{A}_{(\sigma\delta)}$. Im euklidischen Raume bilden also die Mengen, die aus abgeschlossenen Mengen oder Gebieten durch Summen- oder Durchschnittsbildung von Folgen entstehen, immer noch einen verschwindend kleinen Teil des Systems aller Punktmengen.

Wir bezeichnen abgeschlossene Mengen, wie schon öfter, mit F (ensemble fermé), Gebiete mit G. Die F_δ sind wieder abgeschlossen, da ja sogar der Durchschnitt beliebig vieler abgeschlossener Mengen wieder abgeschlossen ist; dagegen sind die F_σ, zu denen u. a. die abzählbaren Mengen gehören, im allgemeinen nicht abgeschlossen. Ebenso sind die G_σ wieder Gebiete, die G_δ nicht. Die F_σ und G_δ sind nächst den F und G selber die wichtigsten und einfachsten Punktmengen; danach folgen die $F_{\sigma\delta}$ und $G_{\delta\sigma}$ usw. Wir wollen alle diese Mengen als Borelsche Mengen bezeichnen. Die F und G, die F_σ und G_δ, die $F_{\sigma\delta}$ und $G_{\delta\sigma}$ sind Komplemente von einander; alle diese Mengen bilden überdies Ringe $\mathfrak{F}, \mathfrak{G}, \mathfrak{F}_\sigma, \mathfrak{G}_\delta, \ldots$.

Wir wollen nun eine beliebige Menge A in ein G_δ einschließen, und zwar zunächst in folgender spezieller Weise: wir wählen eine Folge abnehmender, nach Null konvergierender positiver Zahlen $\varrho_1 > \varrho_2 > \varrho_3 > \ldots$, z. B. $\varrho_n = \frac{1}{n}$, und ordnen jedem Punkt a von A die Umgebung mit dem Mittelpunkt a und Radius ϱ_n zu; die Summe dieser Umgebungen ist ein Gebiet $G_n \geqq A$. Wir behaupten, daß

$$(1) \qquad A_\alpha = \mathfrak{D}(G_1, G_2, \ldots) = G_\delta$$

der Durchschnitt dieser Gebietsfolge ist. In der Tat, ist x ein

α-Punkt von A, so gibt es, für jedes n, einen Punkt a_n von A mit $\overline{xa_n} < \varrho_n$; x gehört also der Umgebung von a_n mit dem Radius ϱ_n, also dem Gebiet G_n an, folglich $A_\alpha \leqq G_\delta$. Umgekehrt, ist y kein Punkt von A_α, und U_y eine Umgebung mit dem Radius ϱ, die keinen Punkt von A enthält, so ist $\overline{ya} \geqq \varrho$ für jedes a, und für $\varrho_n \leqq \varrho$ gehört y nicht mehr zu G_n, also nicht zu G_δ, folglich $G_\delta \leqq A_\alpha$.

Bezeichnen wir die Komplemente von A, G_n, G_δ mit B, F_n, F_σ, so ist

(2) $$B_i = \mathfrak{S}(F_1, F_2, \ldots) = F_\sigma.$$

Die Gleichungen (1), (2), worin A_α jede abgeschlossene Menge und B_i jedes Gebiet bedeuten kann, sagen:

In einem metrischen Raume ist jede abgeschlossene Menge als Durchschnitt einer Folge von Gebieten, jedes Gebiet als Summe einer Folge abgeschlossener Mengen darstellbar.

Oder: jedes F ist ein G_δ, jedes G ein F_σ.

Z. B. ist die abgeschlossene Kreisfläche (Inneres mit Peripherie) vom Radius ϱ der Durchschnitt der Kreisgebiete (Inneres ohne Peripherie) mit den Radien $\frac{2}{1}\varrho$, $\frac{3}{2}\varrho$, $\frac{4}{3}\varrho$, ...; das Kreisgebiet vom Radius ϱ Summe der abgeschlossenen Kreisflächen mit den Radien $\frac{1}{2}\varrho$, $\frac{2}{3}\varrho$, $\frac{3}{4}\varrho$, ...

Eine Differenz zweier abgeschlossener Mengen oder zweier Gebiete oder ein Durchschnitt $\mathfrak{D}(F, G)$ ist sowohl ein F_σ als auch ein G_δ; denn die F und G sind beides, und die F_σ und G_δ bilden ja Ringe. Auch die Mengen des kleinsten Körpers $\mathfrak{F}_\varkappa = \mathfrak{G}_\varkappa$, d. h. die Mengen, die aus abgeschlossenen Mengen oder Gebieten durch endlich wiederholte Differenzbildung entstehen, sind gleichzeitig Mengen F_σ und G_δ. Eine Summe von höchstens abzählbar vielen Differenzen abgeschlossener Mengen ist ein F_σ. In einem metrischen Raume mit abzählbarer dichter Teilmenge trifft dies (§ 4) auf die Mengen $A - A_l$ zu, die von einer beliebigen Menge nach Abzug des letzten Residuums übrigbleiben, insbesondere auf die reduziblen Mengen; da aber auch das Komplement einer reduziblen Menge reduzibel ist, so ist eine reduzible Menge gleichzeitig ein F_σ und ein G_δ. Insbesondere ist eine separierte Menge nicht nur ein F_σ (sogar höchstens abzählbar), sondern auch ein G_δ.

In der Formel (1) haben wir unter G_n die Summe aller Umgebungen von Punkten a mit konstantem Radius ϱ_n (der von a unabhängig ist) verstanden. Läßt man diese Voraussetzung fallen, so braucht der Gebietsdurchschnitt $G_\delta \geqq A$ nicht mehr mit A_α übereinzustimmen. Ordnen wir also jedem Punkte a und jeder natürlichen

Zahl n eine positive Zahl ϱ_{an} zu, verstehen unter U_{an} die Umgebung von a mit dem Radius ϱ_{an} und setzen

$$G_n = \overset{A}{\underset{a}{\mathfrak{S}}} U_{an}, \quad G_\delta = \mathfrak{D}(G_1, G_2, \ldots),$$

so wird man allgemein außer $G_\delta \geqq A$ nichts sagen können. Aber:

Wenn die ϱ_{an} bei festem n eine positive untere Schranke haben, also $\varrho_{an} \geqq \varrho_n > 0$ ist, so bleibt offenbar die Ungleichung

(3) $$G_\delta \geqq A_\alpha, \quad F_\sigma \leqq B_i$$

bestehen (nach wie vor seien B, F_n, F_σ die Komplemente von A, G_n, G_δ).

Wenn andererseits die ϱ_{an} bei festem n eine obere Schranke haben, die mit $\frac{1}{n}$ nach Null konvergiert, also $\varrho_{an} \leqq \varrho_n$ und $\lim \varrho_n = 0$, so bleibt umgekehrt

(4) $$A \leqq G_\delta \leqq A_\alpha, \quad B \geqq F_\sigma \geqq B_i$$

bestehen. Ja, eine geringe Modifikation lehrt, daß ein Punkt y von $E - A_\alpha = B_i$ nicht nur schließlich Punkt von $E - G_n = F_n$, sondern innerer Punkt dieser Menge wird (nämlich z. B. für $\varrho_n \leqq \frac{1}{2}\varrho$, wenn U_y mit dem Radius ϱ keinen Punkt von A enthält). Es ist also in diesem Falle auch noch

$$B_i \leqq \mathfrak{S}(F_{1i}, F_{2i}, \ldots), \quad A_\alpha \geqq \mathfrak{D}(G_{1\alpha}, G_{2\alpha}, \ldots),$$

was in Verbindung mit (4) die Gleichungen

(5) $$A_\alpha = G_{\delta\alpha} = \mathfrak{D}(G_{1\alpha}, G_{2\alpha}, \ldots), \quad B_i = F_{\sigma i} = \mathfrak{S}(F_{1i}, F_{2i}, \ldots)$$

ergibt.

Man kann die Frage stellen, wie man die Radien zu wählen hat, um G_δ möglichst klein, F_σ möglichst groß zu machen. Eine kleinste Menge G_δ über A ist freilich im allgemeinen nicht vorhanden, da der Durchschnitt beliebig vieler G_δ nicht wieder ein G_δ zu sein braucht (man sieht leicht, daß es immer ein G oder $G_\delta \geqq A$ gibt, das einen vorgeschriebenen Punkt von $E - A$ nicht enthält; das kleinste G_δ über A könnte also nur A selbst sein und existiert nur, wenn A selbst ein G_δ ist). Die Frage hat also keinen präzisen Sinn; indessen kann man sich z. B. die Aufgabe stellen, falls A und demgemäß G_δ im Raume dicht ist, trotzdem noch ein ebenfalls dichtes Komplement F_σ zu erzielen.

Bei einer funktionentheoretischen Anwendung, die wir nach E. Borel geben wollen, wird $A = \{a_1, a_2, \ldots, a_p, \ldots\}$ als abzählbar angenommen, dem Punkte a_p eine Umgebung mit dem Radius ϱ_p zugeordnet und deren Summe als G_1 definiert, während G_n die Summe der Umgebungen mit den Radien $\frac{1}{n}\varrho_p$ sein soll. Haben die Radien $\varrho_1, \varrho_2, \ldots$ eine positive untere Schranke, so gilt (3); haben sie eine obere Schranke, so gilt (4) und (5).

In der komplexen Zahlenebene (ohne unendlich fernen Punkt) sei $A = \{a_1, a_2, \ldots\}$ eine abzählbare Menge; wir betrachten die Reihe

$$(\alpha) \qquad \frac{c_1}{x-a_1} + \frac{c_2}{x-a_2} + \ldots = \sum \frac{c_p}{x-a_p}$$

mit einer zunächst beliebigen Folge nichtverschwindender komplexer Koeffizienten c_p, deren Beträge $\gamma_p = |c_p|$ sein mögen. Die Menge $B = E - A$ zerfällt in die Menge C derjenigen Punkte x, wo die Reihe konvergiert und also eine Funktion $f(x)$ darstellt, und in die Menge D derer, wo sie nicht konvergiert (in A verliert die Reihe jeden Sinn). Indem wir die genauere Bestimmung einer solchen Konvergenzmenge C einer späteren Untersuchung vorbehalten, können wir jedenfalls feststellen, daß C ein gewisses F_σ als Teilmenge enthält. Wir wählen eine Folge positiver Zahlen ϱ_p so, daß die Reihe

$$(\beta) \qquad \sum \frac{\gamma_p}{\varrho_p}$$

konvergiert (z. B. $\varrho_p = 2^p \gamma_p$), und bilden mit ihnen als Radien das Gebiet G_1 und die weiteren oben definierten Mengen. Für einen Punkt x von F_n und jedes p ist dann

$$|x - a_p| \geqq \frac{1}{n} \varrho_p, \quad \left| \frac{c_p}{x-a_p} \right| \leqq n \frac{\gamma_p}{\varrho_p};$$

daraus folgt nach bekannten Sätzen, daß die Reihe (α) in F_n absolut und gleichmäßig konvergiert, also eine in F_n stetige, in seinem Innern F_{ni} reguläre Funktion $f(x)$ darstellt. Die Konvergenzmenge C umfaßt also jedenfalls F_σ, und $f(x)$ ist in $\mathfrak{S}(F_{1i}, F_{2i}, \ldots)$ regulär.

Allerdings könnte $F_\sigma = 0$ sein, so daß mit dieser Betrachtung nichts erreicht wäre. Wenn es aber möglich ist, die ϱ_p beschränkt anzunehmen, wofür die Konvergenz der Reihe

$$(\gamma) \qquad \sum \gamma_p$$

notwendig und hinreichend ist, so gelten die Gleichungen (4) (5), und $f(x)$ ist in B_i regulär, also in dem größten Gebiet, in dem es überhaupt definiert ist; dies Gebiet ist von Null verschieden, wenn A_α von E verschieden, A nicht in E dicht ist. Man kann erreichen, daß B_i ein vorgeschriebenes Gebiet G wird, indem man nämlich für A eine in $F = E - G$ dichte abzählbare Menge nimmt (sollte F endlich sein, so reduziert sich die ganze Betrachtung auf eine Trivialität). Daraus, daß die Pole a_p der einzelnen Glieder der Reihe (α) in F dicht liegen, kann man zwar im allgemeinen nicht schließen, daß $f(x)$ über G hinaus nicht fortsetzbar ist; wenn dies aber zutrifft, was allerdings nur unter gewissen, hier nicht zu diskutierenden Bedingungen der Fall ist, so ist damit ein zweiter Beweis des Weierstrassschen Existenzsatzes für eindeutige Funktionen mit vorgeschriebenem Regularitätsgebiet G geliefert (vgl. S. 274).

Wenn die Menge A in E dicht ist, etwa die Menge der rationalen Punkte, so ist $B_i = 0$ und es kommt keine reguläre Funktion zustande. Trotzdem kann F_σ noch von Null verschieden, ja sogar selbst in E dicht sein, so daß die Reihe (α) dann das merkwürdige Verhalten zeigt, daß sowohl die Menge C ihrer Konvergenzstellen als auch die Menge $A + D$ der Stellen, wo sie sinnlos wird oder nicht konvergiert, die Ebene dicht erfüllt. Um das Nichtverschwinden von F_σ zu erzielen, dürfen die ϱ_p jedenfalls nicht mit positiver unterer Schranke gewählt werden, da sonst (3) gelten würde. Eine hinreichende Bedingung erhält man auf Grund späterer Betrachtungen über den Inhalt von Punktmengen (Kap. X), wenn man den Inhalten der Kreisflächen, die das Gebiet G_1 bilden, eine konvergente Summe $\tau = \pi \cdot \Sigma \varrho_p{}^2$ gibt; dann ist das „Flächenmaß" von G_1 höchstens τ, das von G_n höchstens $\frac{1}{n^2}\tau$, das von G_δ Null, also hat G_δ keine inneren Punkte und F_σ ist dicht (sogar von der Mächtigkeit des Kontinuums). Damit aber die ϱ_p so wählbar seien, daß gleichzeitig (β) und $\Sigma \varrho_p{}^2$ konvergiert, ist notwendig und hinreichend, daß

$$(\delta) \qquad \Sigma \gamma_p{}^{\frac{2}{3}}$$

konvergiere. Denn sind α, β zwei positive Zahlen und $\alpha + \beta = 1$, so ist elementar einzusehen, daß $\alpha u + \beta v \geqq u^\alpha v^\beta$ ist für positive u, v; aus der Konvergenz der Reihen Σu_p, Σv_p mit positiven Gliedern folgt also die der Reihe $\Sigma u_p^\alpha v_p^\beta$. Sonach folgt aus der Konvergenz von $\Sigma \frac{\gamma_p}{\varrho_p}$ und $\Sigma \varrho_p{}^2$ die von $\Sigma \left(\frac{\gamma_p}{\varrho_p}\right)^{\frac{2}{3}} (\varrho_p{}^2)^{\frac{1}{3}} = \Sigma \gamma_p{}^{\frac{2}{3}}$. Umgekehrt kann man, wenn (δ) konvergiert, die ϱ_p in der geforderten Weise wählen, z. B. $\varrho_p = \gamma_p{}^{\frac{1}{3}}$.

Die ganze Betrachtung läßt sich genau so im reellen Gebiet anstellen; man wird dann nur $\Sigma \varrho_p$ statt der Quadratsumme konvergent wählen, wofür die Konvergenz von

$$(\varepsilon) \qquad \Sigma \gamma_p{}^{\frac{1}{2}}$$

notwendig und hinreichend ist. Wenn hier A die Menge der rationalen Zahlen ist, so kann man aber die Dichtigkeit von F_σ noch in anderer Weise erzielen, nämlich auf Grund einer von Liouville gefundenen unteren Schranke für die Approximation algebraischer Zahlen durch rationale: allerdings ist hier nicht sicher, daß F_σ eine so hohe Mächtigkeit wie zuvor erlangt.

Ist nämlich x eine (reelle) algebraische Zahl vom Grade $h > 1$, d. h. Wurzel einer irreduziblen Gleichung

$$\varphi(x) = g_0 x^h + g_1 x^{h-1} + \ldots + g_{h-1} x + g_h = 0$$

mit ganzen rationalen Koeffizienten, so gibt es eine nur von x

abhängige positive Zahl M_x derart, daß für jede rationale Zahl $\frac{p}{q}$ (mit positivem Nenner)

$$(\zeta) \qquad \left|x - \frac{p}{q}\right| \geqq 1 : M_x q^h$$

ist. In der Tat ist, wenn $\frac{p}{q}$ zunächst im Intervall $J = (x-1, x+1)$ gewählt wird,

$$\varphi(x) - \varphi\left(\frac{p}{q}\right) = \left(x - \frac{p}{q}\right) \cdot \varphi'(\xi),$$

wo ξ zwischen x und $\frac{p}{q}$, also ebenfalls in J liegt, also wegen $\varphi(x) = 0$

$$\left|\varphi\left(\frac{p}{q}\right)\right| = \left|x - \frac{p}{q}\right| \cdot |\varphi'(\xi)| \leqq \left|x - \frac{p}{q}\right| \cdot M_x,$$

wenn M_x mindestens gleich dem Maximum von $|\varphi'(\xi)|$ im ganzen Intervall J gewählt wird. Ferner ist $\varphi\left(\frac{p}{q}\right) = (g_0 p^h + \dots + g_h q^h) : q^h$ eine nicht verschwindende rationale Zahl mit dem Nenner q^h, also $\left|\varphi\left(\frac{p}{q}\right)\right| \geqq 1 : q^h$, woraus die Ungleichung (ζ) für das Intervall J folgt; wird überdies $M_x \geqq 1$ genommen, so gilt sie auch außerhalb von J, wo $\left|x - \frac{p}{q}\right| > 1$.

Ordnen wir nun z. B. jeder rationalen Zahl, die wir uns in reduzierter Gestalt $\frac{p}{q}$ mit positivem Nenner geschrieben denken, eine Umgebung mit dem Radius $\frac{1}{q^k}$ zu (k eine feste natürliche Zahl > 1); diese Umgebungen sollen das Gebiet G_1, die mit den Radien $\frac{1}{n}\,\frac{1}{q^k}$ das Gebiet G_n bilden. Ist x eine algebraische Zahl vom Grade h $(1 < h \leqq k)$, so ist für $n \geqq M_x$ und jede rationale Zahl $\frac{p}{q}$

$$\left|x - \frac{p}{q}\right| \geqq \frac{1}{M_x q^h} \geqq \frac{1}{n q^h} \geqq \frac{1}{n q^k},$$

x gehört also nicht dem Gebiet G_n, sondern dem Komplement F_n und der Summe F_σ an. Bei dieser Wahl der Radien enthält F_σ also jedenfalls alle algebraischen Zahlen der Grade $2, 3, \dots, k$ und ist eine dichte Menge.

Bei noch rascherer Abnahme der Radien kann F_σ alle algebraische Zahlen (außer den rationalen) enthalten, so daß G_δ nur aus rationalen und transzendenten Zahlen besteht. Wählen wir z. B. für das erste Gebiet G_1 die Umgebung von $\frac{p}{q}$ mit dem Radius ϑ^q, wo $0 < \vartheta < 1$, so ist für eine algebraische Zahl x vom Grade h und jedes $\frac{p}{q}$

$$\left|x - \frac{p}{q}\right| \geqq \frac{1}{M_x q^h} \geqq \frac{1}{n} \vartheta^q,$$

sobald n hinlänglich groß, nämlich $n \geqq M_x \cdot q^h \vartheta^q$ für $q = 1, 2, 3, \ldots$ gewählt wird, und dies ist möglich, weil $q^h \vartheta^q$ mit $\frac{1}{q}$ nach Null konvergiert. Demnach gehört x nicht zu G_n, sondern zu F_n und zu F_σ.

§ 8. Metrische Räume: Bedingungen für kompakte Mengen.

Wir müssen uns jetzt der Frage zuwenden, unter welchen Bedingungen eine Menge kompakt ist.

Wir nannten eine Menge **beschränkt**, wenn die Entfernungen ihrer Punktpaare eine obere Schranke haben, oder, was auf dasselbe hinauskommt, wenn die Menge in einer Kugel liegt (worunter wir die sphärische Umgebung U_x eines Punktes von E mit irgendwelchem Radius ϱ verstehen). Im euklidischen Raume wird sich zeigen, daß kompakte und beschränkte Mengen identisch sind. Daß im allgemeinen die Beschränktheit keine hinreichende Bedingung der Kompaktheit ist, sehen wir sofort, wenn wir unter E die Menge der rationalen Zahlen verstehen; eine Menge A rationaler Zahlen, die nach einer irrationalen Zahl konvergiert, z. B. die der Zahlen $\left(1 + \frac{1}{n}\right)^n$, hat, obwohl beschränkt, in E keinen Häufungspunkt. Die Art, wie man dies durch Einführung der irrationalen Zahlen beseitigt, wird nachher für uns vorbildlich sein.

Jedenfalls aber ist, wie wir gleich sehen werden, für eine kompakte Menge die Bedingung der Beschränktheit und sogar eine noch schärfere notwendig. Wir wollen eine Menge **total beschränkt** nennen, wenn sie für jedes positive ϱ in einer Summe endlich vieler Kugeln vom Radius ϱ enthalten ist. Eine total beschränkte Menge ist gewiß beschränkt; denn ist, für ein bestimmtes ϱ, σ der Maximalabstand der (endlich vielen) Kugelmittelpunkte, so hat jedes Punktpaar der Menge nach dem Dreiecksaxiom eine Entfernung $< 2\varrho + \sigma$.

Jede Teilmenge einer total beschränkten Menge ist total beschränkt. Jede endliche Menge ist total beschränkt.

Für die Anwendung dieses Begriffes ist folgender Satz bequem:

I. *Eine Menge ist dann und nur dann total beschränkt, wenn in jeder unendlichen Teilmenge von ihr Punktepaare mit beliebig kleiner Entfernung vorkommen.*

Unter einem Punktepaar ist hier natürlich ein solches aus zwei verschiedenen Punkten verstanden.

Ist nämlich A eine unendliche, total beschränkte Menge, und ϱ eine beliebig kleine positive Zahl, so muß mindestens eine von den

endlich vielen Kugeln vom Radius $\frac{\varrho}{2}$, die A einschließen, zwei (sogar unendlich viele) verschiedene Punkte von A enthalten, und deren Entfernung ist $< \varrho$. In einer unendlichen, total beschränkten Menge gibt es also Punktepaare mit beliebig kleiner Entfernung, also auch in jeder unendlichen Teilmenge einer total beschränkten Menge. Wenn umgekehrt jede unendliche Teilmenge von A Punktpaare mit beliebig kleiner Entfernung enthält, so sei a_1 ein beliebiger Punkt von A, a_2 ein weiterer mit $\overline{a_1 a_2} \geqq \varrho$, a_3 ein dritter mit $\overline{a_1 a_3} \geqq \varrho$, $\overline{a_2 a_3} \geqq \varrho$ usw. Dies Verfahren muß einmal abbrechen, auf Grund der Voraussetzung, d. h. für ein bestimmtes n gibt es keinen Punkt a_{n+1} mehr, der von allen vorigen eine Entfernung $\geqq \varrho$ hätte. Dann ist aber A in der Summe der n Kugeln mit den Mittelpunkten $a_1, a_2, \ldots, a_n$ und den Radien ϱ enthalten; da dies für jedes ϱ durchführbar ist, so ist A total beschränkt.

Wir sahen, daß eine total beschränkte Menge auch beschränkt ist. Im euklidischen Raume ist umgekehrt jede schlechthin beschränkte Menge auch total beschränkt; das folgt (wenn wir statt der Kugeln mit Würfeln operieren, denen wir nebenbei ihre Begrenzung zurechnen wollen) höchst einfach aus der Möglichkeit, einen Würfel in eine endliche Anzahl von Teilwürfeln zu zerlegen, in denen der Maximalabstand zweier Punkte beliebig klein wird. In einem allgemeinen metrischen Raume ist die totale Beschränktheit aber keineswegs eine Folge der einfachen. Das kann man schon daraus schließen, daß man jeden metrischen Raum durch eine veränderte Definition der Entfernung in einen beschränkten verwandeln kann, indem man z. B. die Entfernung ϱ durch $\varrho' = \frac{\varrho}{1+\varrho} < 1$ ersetzt.[1] Die ganzzahligen Punkte einer Geraden haben jetzt paarweise Entfernungen $\geqq \frac{1}{2}$ und bilden also nach I keine total beschränkte Menge, wohl aber eine beschränkte. Ein anderes Beispiel geben die Punkte

$$(1, 0, 0, 0, \ldots) \quad (0, 1, 0, 0, \ldots) \quad (0, 0, 1, 0, \ldots) \quad \ldots$$

eines Hilbertschen Raumes (§ 5, IV), die paarweise die Entfernung $\sqrt{2}$ haben.

Eine Menge mit einem Häufungspunkt enthält gewiß Punktpaare mit beliebig kleiner Entfernung (gleichviel ob der Häufungspunkt ihr angehört oder nicht); daraus folgt:

II. Jede kompakte Menge ist total beschränkt.

Wir haben dies auch schon beim Beweise von § 3, X eingesehen.

[1] Auch diese Entfernungen erfüllen das Dreiecksaxiom; aus $\varrho + \sigma \geqq \tau$ folgt nämlich $\varrho' + \sigma' \geqq \tau'$.

Daß II nicht umkehrbar, also die totale Beschränktheit nur eine notwendige, keine hinreichende Bedingung für Kompaktheit ist, lehrt immer noch die Menge der rationalen Zahlen (als Raum E gedacht). Wir werden aber jetzt zeigen, daß eine Annahme, die der Einführung der irrationalen Zahlen analog ist, den Satz II umkehrbar macht, und daß man diese Annahme, falls sie von vornherein nicht erfüllt ist, durch nachträgliche „Vervollständigung" des Raumes verwirklichen kann. Mehr läßt sich offenbar nicht erzwingen, ohne die Metrik des Raumes zu zerstören: eine nicht total beschränkte Menge behält nach I diesen Charakter, solange man die Entfernungen ihrer Punktpaare nicht ändert, und ist in keinem metrischen Raume kompakt. (Die euklidische Ebene E wird durch Hinzufügung eines unendlich fernen Punktes nach § 5, II kompakt, aber $E + \{\infty\}$ ist kein metrischer Raum.)

Unter den total beschränkten Mengen zeichnen wir gewisse als Fundamentalmengen aus: eine Fundamentalmenge heiße eine unendliche Menge, von der, für jedes positive ϱ, fast alle Punkte in einer Kugel vom Radius ϱ liegen. Daß eine Fundamentalmenge total beschränkt ist, bedarf keines Beweises. Eine konvergente Menge ist eine Fundamentalmenge, denn eine Kugel mit ihrem Limes als Mittelpunkt enthält bei beliebigem Radius fast alle Punkte der Menge. Umgekehrt können wir zunächst nur behaupten, daß eine Fundamentalmenge A entweder konvergent oder divergent ist. Denn ist sie nicht divergent, so sei x einer ihrer Häufungspunkte. Für ein bestimmtes ϱ sei y der Mittelpunkt einer Kugel vom Radius ϱ, die fast alle Punkte a der Menge A enthält. Da nun, für unendlich viele a, $\overline{xa} < \varrho$ und für fast alle $\overline{ya} < \varrho$ ist, also noch für unendlich viele beides stattfindet, so muß $\overline{xy} < 2\varrho$ sein; dann ist aber, für fast alle a, $\overline{xa} < 3\varrho$, also ist $x = \lim A$.

III. Jede unendliche, total beschränkte Menge enthält eine Fundamentalmenge als Teilmenge.

A sei unendlich und total beschränkt; $\varrho_1 > \varrho_2 > \dots$ eine abnehmende, nach 0 konvergente Folge positiver Zahlen. Es gibt, wie wir sahen, eine Kugel U_1 vom Radius ϱ_1, die eine unendliche Teilmenge A_1 von A enthält; a_1 sei ein Punkt von A_1. Ebenso gibt es weiter, da $A_1 - \{a_1\}$ auch noch unendlich und total beschränkt ist, eine Kugel U_2 vom Radius ϱ_2, die eine unendliche Teilmenge A_2 der letztgenannten Menge enthält; ist wieder a_2 ein Punkt von A_2, so gibt es eine Kugel U_3 vom Radius ϱ_3, die eine unendliche Teilmenge A_3 von $A_2 - \{a_2\}$ enthält usw. Nach dieser Konstruktion ist also U_n eine Kugel vom Radius ϱ_n, die eine unendliche Teilmenge A_n

von A enthält; ferner $A_1 > A_2 > A_3 > \ldots$ und a_n ein Punkt von $A_n - A_{n+1}$. Die Punkte a_n sind paarweise verschieden und bilden eine Fundamentalmenge. Denn ist ϱ beliebig und n so groß, daß $\varrho_n \leqq \varrho$, so sind fast alle Punkte dieser Menge, nämlich $a_n, a_{n+1}, a_{n+2}, \ldots$, in U_n, also auch in einer Kugel vom Radius ϱ enthalten.

Soll nun der Satz II umkehrbar, also jede total beschränkte Menge kompakt sein, so muß insbesondere jede Fundamentalmenge kompakt, also konvergent sein. Umgekehrt, ist jede Fundamentalmenge konvergent, so ist jede total beschränkte Menge kompakt; denn jede unendliche Teilmenge A von ihr enthält eine Fundamentalmenge, und deren Limes ist ein Häufungspunkt von A. Wir sehen also:

IV. Damit jede total beschränkte Menge kompakt sei, ist die Konvergenz jeder Fundamentalmenge notwendig und hinreichend.

In einer gebräuchlicheren Form läßt sich dies auch mit Hilfe von Punktfolgen ausdrücken. Wir nennen $(a_1, a_2, \ldots)$ eine Fundamentalfolge, wenn für jedes positive ϱ fast alle Punkte in einer Kugel vom Radius ϱ liegen, wenn also für jedes ϱ ein Punkt x und eine Zahl n existiert derart, daß $a_n, a_{n+1}, \ldots$ von x eine Entfernung $< \varrho$ haben. Alle diese Punkte haben dann voneinander eine Entfernung $< 2\varrho$, und wir können eine Fundamentalfolge also auch durch die zweite Bedingung definieren[1]: für jedes positive ϱ soll eine Zahl n existieren so, daß $\overline{a_p a_q} < \varrho$ für $p \geqq n, q \geqq n$. Wie oben sieht man, daß eine konvergente Folge eine Fundamentalfolge ist, und umgekehrt eine Fundamentalfolge nur konvergent oder divergent sein kann. Eine leichte Überlegung formt IV in den Satz um:

V. Damit jede total beschränkte Menge kompakt sei, ist die Konvergenz jeder Fundamentalfolge notwendig und hinreichend.

Denn eine divergente Fundamentalmenge liefert in ihren abzählbaren Teilmengen divergente Fundamentalfolgen, und eine divergente Fundamentalfolge liefert als Menge ihrer verschiedenen Punkte eine divergente Fundamentalmenge. Konvergenz aller Fundamentalmengen und Konvergenz aller Fundamentalfolgen sind also gleichbedeutend.

---

[1] Oder: für jedes ϱ soll ein n existieren derart, daß $\overline{a_n a_p} < \varrho$ für $p > n$. Oder: für jede Folge natürlicher Zahlen $p_1 < p_2 < \ldots$ soll $\lim \overline{a_n a_{p_n}} = 0$ sein. Wenn die a_n eine Fundamentalfolge bilden und $\lim \overline{a_n b_n} = 0$, so bilden auch die b_n eine Fundamentalfolge.

Wir haben (Kap. VII, § 9) gesehen, daß im eindimensionalen euklidischen Raume oder in der Menge T der reellen Zahlen jede beschränkte, also erst recht jede total beschränkte Menge kompakt ist. Daraus folgt, daß in T jede Fundamentalfolge konvergiert:

VI (Konvergenztheorem von Cauchy). Eine Folge reeller Zahlen konvergiert dann und nur dann, wenn sie eine Fundamentalfolge ist.

Dies ist ja in der Tat das von Cauchy herrührende, sogenannte allgemeine Konvergenzkriterium.

Wir wollen einen metrischen Raum vollständig nennen, wenn jede Fundamentalfolge konvergiert.[1] In einem vollständigen Raume sind also kompakte Mengen und total beschränkte Mengen identisch.

Wir zeigen schließlich noch, wie man einen beliebigen metrischen Raum E stets zu einem vollständigen erweitern oder, präziser, auf eine Teilmenge eines vollständigen Raumes E umkehrbar eindeutig und entfernungstreu abbilden kann. Das Verfahren ist genau der Methode von G. Cantor und Ch. Méray nachgebildet, die irrationalen Zahlen durch Fundamentalfolgen rationaler Zahlen zu definieren[2], und besteht einfach darin, die Fundamentalfolgen $\alpha = (a_1, a_2, \ldots)$ aus Punkten des Raumes E ihrerseits als Elemente eines zweiten Raumes E anzusehen.

Sind $\alpha = (a_1, a_2, \ldots)$, $\beta = (b_1, b_2, \ldots)$ zwei Fundamentalfolgen, so bilden die Entfernungen $\varrho_n = \overline{a_n b_n}$ entsprechender Punkte eine Fundamentalfolge reeller Zahlen. In der Tat: bei beliebigem positivem ε gibt es eine Zahl n derart, daß, für $p, q \geqq n$, $\overline{a_p a_q}$ und $\overline{b_p b_q} < \varepsilon$ sind. Dann ist

$$\overline{a_p b_p} \leqq \overline{a_p a_q} + \overline{a_q b_q} + \overline{b_q b_p} < \overline{a_q b_q} + 2\varepsilon,$$

also $\varrho_p - \varrho_q < 2\varepsilon$ und, wenn man q und p vertauscht, $|\varrho_p - \varrho_q| < 2\varepsilon$, woraus das Gesagte hervorgeht. Wegen des Cauchyschen Theorems existiert also $\lim \varrho_n$, und diese Zahl wird als Entfernung

$$\overline{\alpha\beta} = \lim \overline{a_n b_n}$$

der beiden Fundamentalfolgen definiert. Sie erfüllt das Dreiecksaxiom, da aus $\overline{a_n b_n} + \overline{b_n c_n} \geqq \overline{a_n c_n}$ durch Grenzübergang $\overline{\alpha\beta} + \overline{\beta\gamma} \geqq \overline{\alpha\gamma}$ folgt; um das Koinzidenzaxiom gültig zu machen, müssen wir ähnlich wie in einem früheren Falle (§ 6) für zwei Fundamentalfolgen

---

[1] M. Fréchet sagt: der Raum „gestattet eine Verallgemeinerung des Theorems von Cauchy".

[2] Natürlich sind, solange nur die rationalen Zahlen existieren und die irrationalen erst geschaffen werden sollen, gewisse Vorsichtsmaßregeln notwendig, die uns auf unserem Standpunkt erspart bleiben.

mit verschwindender Entfernung eine Verabredung treffen, sei es, daß wir nur eine von ihnen in E aufnehmen oder für $\overline{\alpha\beta}=0$ einfach $\alpha=\beta$ definieren mit erweiterter Bedeutung des Gleichheitszeichens.

Unter den Fundamentalfolgen sind die **konstanten** $(a, a, \ldots)$ aus lauter gleichen Punkten ausgezeichnet; ihre Menge ist auf E umkehrbar eindeutig und entfernungstreu bezogen, da die Entfernung der Folgen $(a, a, \ldots)$ und $(b, b, \ldots)$ gleich $\overline{ab}$ ist. Es wird daher erlaubt und im Interesse einfacherer Bezeichnung erwünscht sein, die konstante Folge $(a, a, \ldots)$ direkt mit a zu identifizieren, so daß E Teilmenge von E wird. E ist in E dicht. Denn zu jeder Fundamentalfolge $\alpha=(a_1, a_2, \ldots)$ gibt es bei beliebigem ϱ einen Punkt x derart, daß, für hinlänglich großes n, $\overline{xa_n}<\varrho$ wird; daraus folgt $\lim\overline{xa_n}\leqq\varrho$ oder $\overline{x\alpha}\leqq\varrho$, d. h. die Fundamentalfolge α und die konstante Folge x haben eine Entfernung $\leqq\varrho$. Weiter ist aber, wenn die konstanten Folgen $a_n=(a_n, a_n, \ldots)$ in Betracht gezogen werden, $\overline{\alpha a_n}\leqq\overline{\alpha x}+\overline{xa_n}<2\varrho$ für hinlänglich großes n, also $\lim\overline{\alpha a_n}=0$. Nennen wir die Folgen von Elementen aus E der Deutlichkeit wegen Sequenzen, so können wir also sagen: bilden die a_n in E eine Fundamentalfolge α, so bilden die konstanten Folgen a_n in E eine konvergente Sequenz mit dem Limes α.

Nunmehr ist die Vollständigkeit von E leicht zu beweisen. Wir sahen (S. 314 Anm.), daß von zwei Sequenzen $\alpha_1, \alpha_2, \ldots$ und $\beta_1, \beta_2, \ldots$ mit $\lim\overline{\alpha_n\beta_n}=0$ die eine zugleich mit der andern eine Fundamentalsequenz ist. Ist nun $\alpha_1, \alpha_2, \ldots$ eine Fundamentalsequenz, so kann man wegen der Dichtigkeit von E in E zu jedem α_n eine konstante Folge x_n mit $\overline{\alpha_n x_n}<\frac{1}{n}$ oder $\lim\overline{\alpha_n x_n}=0$ bestimmen. Dann bilden die x_n ebenfalls eine Fundamentalsequenz, also in E eine Fundamentalfolge[1] $\xi=(x_1, x_2, \ldots)$, und es ist $\lim\overline{\xi x_n}=0$, also auch $\lim\overline{\xi\alpha_n}=0$, d. h. die Fundamentalsequenz $\alpha_1, \alpha_2, \ldots$ ist konvergent mit dem Limes ξ.

Der Übergang von E zu E ist das metrische Analogon zur Ausfüllung der Lücken geordneter Mengen (Kap. IV, § 5); für diese gibt die **Dedekind**sche, für jenen die **Cantor**sche Theorie der Irrationalzahlen das Vorbild.

Die wichtigste Klasse vollständiger Räume ist die der **euklidischen**. Bilden in E_n die Punkte $a_1, a_2, \ldots$ eine Fundamentalfolge und sind $a_{p1}, a_{p2}, \ldots, a_{pn}$ die Koordinaten von a_p, so folgt aus

[1] Wegen der Entfernungstreue: man nehme die zweite Definition der Fundamentalfolge.

der Definition der Entfernung, daß auch die ersten Koordinaten $a_{11}, a_{21}, \ldots$ eine Fundamentalfolge bilden (weil $|a_{p1} - a_{q1}| \leqq \overline{a_p a_q}$) und also einen Limes x_1 haben; das gleiche gilt für die übrigen Koordinaten, die nach Limites $x_2, \ldots, x_n$ konvergieren. Für den Punkt $x = (x_1, x_2, \ldots, x_n)$ folgt dann aber $\lim \overline{x a_p} = 0$, also $x = \lim a_p$. Also ist E_n ein vollständiger Raum; danach sind kompakte und total beschränkte Mengen, überdies (S. 312) total beschränkte und schlechthin beschränkte Mengen identisch; also sind kompakte und beschränkte Mengen identisch.

Auch der Hilbertsche Raum (§ 5, IV) ist vollständig. Bilden die Punkte $a_1, a_2, \ldots$ eine Fundamentalfolge und ist

$$a_p = (a_{p1}, a_{p2} \ldots, a_{pn}, \ldots),$$

so folgt wie oben, daß $\lim\limits_p a_{pn} = x_n$ existiert. Zu jedem positiven ϱ ist eine Zahl m angebbar so, daß für $p \geqq m$, $q \geqq m$

$$\overline{a_p a_q} < \varrho, \quad \sum_n (a_{pn} - a_{qn})^2 < \varrho^2.$$

Für jede natürliche Zahl N ist dann auch

$$\sum_1^N {}_n (a_{pn} - a_{qn})^2 < \varrho^2.$$

Daraus folgt für $\lim \frac{1}{q} = 0$

$$\sum_1^N {}_n (a_{pn} - x_n)^2 \leqq \varrho^2,$$

daraus für $\lim \frac{1}{N} = 0$

$$\sum_n (a_{pn} - x_n)^2 \leqq \varrho^2.$$

Also hat die Zahlenfolge $x = (x_1, x_2, \ldots)$ von a_p eine Entfernung (d. h. x gehört dem Hilbertschen Raume an, $\sum x_n{}^2$ konvergiert) und es ist $\overline{x a_p} \leqq \varrho$ für $p \geqq m$, also $\lim \overline{x a_p} = 0$, die Fundamentalfolge konvergiert nach x.

Der „Raum“ der stetigen Funktionen (§ 5, V) ist bei der ersten Entfernungsdefinition vollständig; denn eine Fundamentalfolge konvergiert gleichmäßig gegen einen Limes, der mithin wieder eine stetige Funktion ist (die Ausführung können wir dem Leser überlassen).

Jede abgeschlossene Menge eines vollständigen Raumes ist wieder ein vollständiger Raum. Ein kompakter metrischer Raum ist a fortiori vollständig; jede kompakte abgeschlossene Menge eines metrischen Raumes ist ein kompakter metrischer Raum, also vollständig. Natürlich ist dieser Übergang von einem Raum zu seiner Teilmenge mit unveränderten Entfernungen gedacht.

§ 9. Vollständige Räume.

Die wesentlichste Eigenschaft eines vollständigen Raumes ist, daß er über die Mächtigkeit gewisser Mengen eine Aussage gestattet. Es gilt zunächst der Satz:

I. In einem vollständigen Raume hat eine absteigende Folge beschränkter, abgeschlossener, von Null verschiedener Mengen $A_1 \geqq A_2 \geqq \ldots$, deren Breite nach Null konvergiert, stets einen und nur einen gemeinsamen Punkt.

Unter der Breite einer beschränkten Menge verstanden wir die obere Schranke der Entfernungen ihrer Punktpaare. Daß $\mathfrak{D}(A_1, A_2, \ldots)$ nicht zwei verschiedene Punkte haben kann, wenn die Breite ϱ_n von A_n nach 0 konvergiert, ist selbstverständlich. Andererseits sei a_n ein Punkt von A_n; die Punkte $a_1, a_2, \ldots$ bilden dann eine Fundamentalfolge, da die zu A_n gehörigen Punkte $a_n, a_{n+1}, \ldots$ alle von a_n einen Abstand $\leqq \varrho_n$ haben, den man beliebig klein machen kann. Also existiert $x = \lim a_n$; dies ist ein α-Punkt für jedes A_n, also ein Punkt von A_n selbst und demnach von $\mathfrak{D}(A_1, A_2, \ldots)$.

Der Satz I hat eine große Ähnlichkeit mit dem Cantorschen Durchschnittssatz (Kap. VII, § 4, I); er ist aber zugleich allgemeiner als dieser, da er von beschränkten Mengen handelt, die ja in einem vollständigen Raume nicht kompakt zu sein brauchen[1], und spezieller, insofern er die Bedingung der nach Null konvergierenden Breite hinzufügt.

Wir werden den Satz insbesondere auf abgeschlossene Kugeln V_x anwenden: darunter verstehen wir die Menge der Punkte y, die von einem Punkt x eine Entfernung $\leqq \varrho$ haben, während U_x nach wie vor die Menge der Punkte mit $\overline{xy} < \varrho$ bedeutet. Daß $E - V_x$ ein Gebiet, also V_x abgeschlossen ist, bedarf keines Beweises. Die durch gleichen Mittelpunkt und Radius verbundenen U_x, V_x mögen entsprechend heißen; man beachte übrigens, daß im allgemeinen nicht, wie im euklidischen Raume, U_x mit der Menge V_{xi} der inneren Punkte von V_x und V_x mit $U_{x\alpha}$ identisch zu sein braucht. Ist $\varrho' < \varrho < \varrho''$, so ist $V_x' \leqq U_x \leqq V_x \leqq U_x''$. Jeder Punkt x eines Gebietes hat auch eine dem Gebiet angehörige abgeschlossene Umgebung V_x.

II (Satz von Cantor). In einem vollständigen Raume ist eine abgeschlossene Menge mit nichtverschwindendem Kern mindestens von der Mächtigkeit des Kontinuums.

---

[1] Das Beispiel S. 312 gab eine beschränkte, aber nicht total beschränkte, also nicht kompakte Menge im vollständigen Hilbertschen Raume.

III (Satz von Young). In einem vollständigen Raume ist eine Menge G_δ (Durchschnitt einer Gebietsfolge) mit nichtverschwindendem Kern mindestens von der Mächtigkeit des Kontinuums.

Wir beweisen beide Sätze gemeinschaftlich; überdies ist II ein Spezialfall von III, da jede abgeschlossene Menge eines metrischen Raumes ein G_δ ist (§ 7).

Zunächst sei A eine nichtverschwindende insichdichte Menge, a_1 und a_2 zwei ihrer Punkte, die wir mit abgeschlossenen Kugeln V_1 und V_2 umgeben, die keinen Punkt gemeinsam haben. In dem zu V_1 entsprechenden Kugelinnern U_1 liegen, da a_1 Häufungspunkt von A ist, unendlich viele Punkte dieser Menge; zwei davon, a_{11} und a_{12}, umgeben wir mit abgeschlossenen Kugeln V_{11} und V_{12}, die in V_1 liegen und keinen gemeinsamen Punkt haben. Ebenso verfahren wir mit V_2: zwei Punkte a_{21} und a_{22}, die zu A gehören und in U_2 liegen, umgeben wir mit abgeschlossenen Kugeln V_{21} und V_{22}, die zueinander fremd sind und in V_2 liegen. So fahren wir fort: da a_{11} Häufungspunkt von A ist und in dem zu V_{11} entsprechenden U_{11} unendlich viele Punkte von A liegen, so können wir zwei von diesen Punkten, a_{111} und a_{112}, mit abgeschlossenen, in V_{11} liegenden und zueinander fremden Kugeln V_{111} und V_{112} umgeben usw. Für jede aus den Ziffern 1, 2 gebildete Ziffernfolge $(p, q, r, \ldots)$ erhalten wir so eine Folge abgeschlossener Kugeln

$$V_p \geqq V_{pq} \geqq V_{pqr} \geqq \ldots,$$

deren Mittelpunkte $a_p, a_{pq}, a_{pqr}, \ldots$ Punkte von A sind, und wobei V_1 und V_2, V_{p1} und V_{p2}, V_{pq1} und $V_{pq2}, \ldots$ keinen gemeinsamen Punkt haben.

Bei dieser Konstruktion hindert nichts, die Radien von V_1 und V_2 etwa kleiner als 1, die von $V_{11}, V_{12}, V_{21}, V_{22}$ kleiner als $\frac{1}{2}$, die der acht nächstfolgenden Kugeln kleiner als $\frac{1}{3}$ zu wählen usf. Für jede der obengenannten Kugelfolgen konvergieren dann die Radien oder die Breiten der Kugeln nach Null; demnach bestimmt jede solche Folge einen Durchschnitt

$$\{x\} = \mathfrak{D}(V_p, V_{pq}, V_{pqr}, \ldots),$$

der sich auf einen einzigen Punkt reduziert, und wegen der Fremdheit von V_1 und V_2 usw. gehören zu verschiedenen Zahlenfolgen $(p, q, r, \ldots)$ verschiedene Punkte x. Die Menge dieser Durchschnittspunkte ist also von der Mächtigkeit $2^{\aleph_0} = \aleph$ des Kontinuums.

Hiernach vollendet sich der Beweis der fraglichen Sätze folgendermaßen:

Ist A in einer abgeschlossenen Menge F enthalten, so ist jedes x, als Limes der Fundamentalfolge $(a_p, a_{pq}, a_{pqr}, \ldots)$, ein α-Punkt

von A, also Punkt von F; F hat mindestens die Mächtigkeit des Kontinuums.

Ist A in einer Menge $G_\delta = \mathfrak{D}(G_1, G_2, \ldots)$ enthalten, so unterwerfen wir die V noch einer weiteren Bedingung: da a_1 und a_2 Punkte des Gebiets G_1 sind, können wir V_1 und V_2 in G_1 liegend annehmen, ebenso V_{11}, V_{12}, V_{21}, V_{22} in G_2, die acht nächstfolgenden Kugeln in G_3 usw. Dann ist

$$\mathfrak{D}(V_p, V_{pq}, V_{pqr}, \ldots) \leqq \mathfrak{D}(G_1, G_2, G_3, \ldots),$$

also jedes x Punkt von G_δ, G_δ hat mindestens die Mächtigkeit $\aleph$.

Damit sind die Sätze bewiesen; wir knüpfen die weiteren Betrachtungen an den allgemeineren Satz III an. Es folgt sogleich noch mehr: auch V_1 enthält mindestens $2^{\aleph}$ Punkte x, und da V_1 einen beliebig kleinen Radius haben darf, so ist a_1 Verdichtungspunkt von G_δ. Hierbei ist a_1 ein beliebiger Punkt von A, A eine insichdichte Teilmenge von G_δ. Der Kern der Menge G_δ ist also in der Menge ihrer Verdichtungspunkte enthalten.

Von besonderer Bedeutung werden diese Resultate in einem vollständigen Raum mit abzählbarer dichter Teilmenge, wo also die Folgen des zweiten Abzählbarkeitsaxioms (§ 3) hinzutreten. Wir sahen hier, daß für jede Menge A die Menge

$$A_v = \mathfrak{D}(A, A_\gamma)$$

der zu ihr gehörigen Verdichtungspunkte insichdicht, also im Kern A_k enthalten ist ($A_v \leqq A_k$). Andererseits, ist A eine Menge G_δ, so hatten wir soeben gefunden, daß $A_k \leqq A_\gamma$ oder $A_k \leqq A_v$; beides zusammen gibt $A_k = A_v$. Für eine abgeschlossene Menge A ist insbesondere $A_k = A_\gamma$, und für eine perfekte Menge, die ja ihr eigener Kern ist, $A = A_\gamma$. (Daß umgekehrt eine Menge A, für die $A = A_\gamma$, perfekt ist, konnte schon im vorigen Kapitel bewiesen werden, denn dann ist A abgeschlossen und insichdicht, weil A_γ abgeschlossen und $A \leqq A_\beta$ ist.) Ferner ist $A - A_v$ höchstens abzählbar und endlich E selbst, also jede Punktmenge höchstens von der Mächtigkeit des Kontinuums. Das gibt folgendes Resultat:

IV. In einem vollständigen Raum mit abzählbarer dichter Teilmenge ist jede Menge A, die ein G_δ (Durchschnitt einer Folge von Gebieten) ist, von der Mächtigkeit des Kontinuums, falls sie einen nichtverschwindenden Kern hat, andernfalls höchstens abzählbar. Der Kern A_k ist mit der Menge $A_v = \mathfrak{D}(A, A_\gamma)$ der zu A gehörigen Verdichtungspunkte identisch; für eine abgeschlossene Menge ist $A_k = A_\gamma$, für eine perfekte Menge $A = A_\gamma$.

Anwendungen. Eine perfekte Menge, die nicht Null ist, ist von der Mächtigkeit des Kontinuums.

Eine Menge kann höchstens auf eine Weise in einen perfekten und einen höchstens abzählbaren Bestandteil gespalten werden. Denn soll $A = P + Q$, P perfekt, Q höchstens abzählbar sein, so ist $P_\gamma = P$, $Q_\gamma = 0$, also $A_\gamma = \mathfrak{S}(P_\gamma, Q_\gamma) = P$; der perfekte Bestandteil ist A_γ. Die Zerlegung ist also dann und nur dann möglich, wenn $A \geqq A_\gamma$ (eine solche Menge braucht weder abgeschlossen noch insichdicht zu sein).

Eine abzählbare Menge, die einen nichtverschwindenden Kern hat (z. B. die der rationalen Punkte eines euklidischen Raumes), ist kein G_δ, wohl aber wie jede abzählbare Menge ein F_σ. Das Komplement einer solchen Menge (z. B. die der irrationalen Punkte) ist kein F_σ, wohl aber ein G_δ. Da jede separierte Menge ein G_δ ist (S. 306), so ist, damit eine abzählbare Menge ein G_δ sei, notwendig und hinreichend, daß sie separiert sei.

Wüßte man für alle Mengen eines euklidischen Raumes, was man nach IV von den abgeschlossenen Mengen F oder den Mengen G_δ weiß, daß sie nämlich endlich, abzählbar oder von der Mächtigkeit $\aleph$ sind, so wäre $\aleph$ die nächste Mächtigkeit über $\aleph_0$ und damit die Kontinuumfrage im Sinne der Cantorschen Vermutung $\aleph = \aleph_1$ entschieden. Um aber einzusehen, wie weit man noch von diesem Ziel entfernt ist, genügt es sich zu erinnern, daß das System der Mengen F oder G_δ nur einen verschwindend kleinen Teil des Systems aller Punktmengen bildet (S. 305).

Offenbar ist auch jede Menge $F_\sigma = \mathfrak{S}(F_1, F_2, \ldots)$ endlich, abzählbar oder von der Mächtigkeit $\aleph$, denn entweder sind alle F_n höchstens abzählbar oder mindestens eins ist von der Mächtigkeit $\aleph$; dasselbe gilt von den Mengen $G_{\delta\sigma}$. Für diese Mengen entscheidet aber nicht mehr das Verschwinden oder Nichtverschwinden des Kerns über die Mächtigkeit; eine Menge F_σ mit nichtverschwindendem Kern kann abzählbar sein (z. B. die Menge der rationalen Punkte).

Jede kompakte unendliche Menge A eines metrischen Raumes kann als Teilmenge des kompakten, also vollständigen Raumes A_α mit abzählbarer dichter Teilmenge (§ 3, X) aufgefaßt werden; ist sie also abgeschlossen oder ein G_δ, so ist sie abzählbar oder von der Mächtigkeit $\aleph$, je nachdem ihr Kern verschwindet oder nicht. Die Komponenten einer kompakten abgeschlossenen metrischen Menge bilden nach § 6, XI wieder eine solche, im Sinne der Distanzen; eine kompakte abgeschlossene Menge hat also entweder höchstens abzählbar viele oder $\aleph$ Komponenten, je nachdem der Kern der Komponentenmenge verschwindet oder nicht. (Der Bestimmung des Kerns sind die Distanzen zugrunde zu legen, die, wie erinnerlich, in gewisser Weise durch die unteren Entfernungen vertreten werden können.)

Bei dem Beweise der Sätze II, III haben wir eine (spezielle) Menge X folgender Art verwendet: es seien $A_1, A_2, A_{11}, A_{12}, A_{21}, A_{22}, \ldots$

nichtverschwindende abgeschlossene beschränkte Mengen; für jede aus den Ziffern 1, 2 gebildete Folge $(p, q, r, \ldots)$ sei

$$A_p > A_{pq} > A_{pqr} > \ldots$$

und die Breite dieser Mengen konvergiere nach 0; endlich sei A_1 fremd zu A_2, A_{p1} fremd zu A_{p2}, A_{pq1} fremd zu A_{pq2} usw. Dann soll X die Menge der Durchschnittspunkte oder die Summe der Durchschnitte

$$\{x\} = \mathfrak{D}(A_p, A_{pq}, A_{pqr}, \ldots)$$

dieser absteigenden Folgen sein, wofür man auch den Durchschnitt der Summen

$$X = \mathfrak{D}(\Sigma A_p, \Sigma A_{pq}, \Sigma A_{pqr}, \ldots)$$

setzen kann. Eine solche Menge X, die wir zur Abkürzung eine dyadische Menge nennen wollen, ist abgeschlossen, offenbar auch perfekt, und von der Mächtigkeit des Kontinuums. Ferner ist sie punkthaft, wenn wir nämlich unter einer punkthaften Menge eine solche verstehen, deren sämtliche Komponenten aus einzelnen Punkten bestehen. Denn eine zusammenhängende Teilmenge von X ist auch eine solche von ΣA_p, ΣA_{pq}, ... und muß dabei jedesmal ganz in einem Summanden liegen, also einer einzigen Folge A_p, A_{pq}, A_{pqr}, ... angehören.

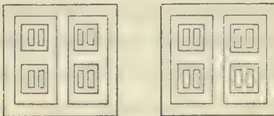


Fig. 11.

Sind x, x' zwei verschiedene Punkte von X und $(p, q, r, \ldots)$, $(p', q', r', \ldots)$ die entsprechenden verschiedenen Ziffernfolgen, so ist hiermit eine natürliche Zahl $k = k(x, x')$ bestimmt derart, daß die Ziffernfolgen in den ersten $k-1$ Ziffern übereinstimmen, in der k^{ten} nicht. Wenn wir x und seine Ziffernfolge $(p, q, r, \ldots)$ festhalten und setzen:

$$\begin{aligned}
\delta_1 &= \delta(A_1, A_2), & d_1 &= d(A_1, A_2) \\
\delta_2 &= \delta(A_{p1}, A_{p2}), & d_2 &= d(A_{p1}, A_{p2}) \leqq d(A_p) \\
\delta_3 &= \delta(A_{pq1}, A_{pq2}), & d_3 &= d(A_{pq1}, A_{pq2}) \leqq d(A_{pq}) \quad \text{usw.},
\end{aligned}$$

wo wie früher $\delta(A, B)$, $d(A, B)$ die untere und obere Entfernung der Mengen A, B und $d(A)$ die Breite von A bedeutet, so ist offenbar

$$\text{für } k(x, x') = k\colon \quad \delta_k \leqq \overline{xx'} \leqq d_k.$$

Hieraus ergibt sich unmittelbar: wenn x_n eine Folge von x verschiedener Punkte aus X durchläuft und $k(x, x_n) = k_n$ gesetzt wird, so folgt aus $\lim \overline{xx_n} = 0 : \lim \delta_{k_n} = 0$, also $\lim k_n = \infty$ (weil die δ_k positiv sind), und umgekehrt aus $\lim k_n = \infty : \lim d_{k_n} = 0$, also $\lim \overline{xx_n} = 0$ (weil die Breiten von A_p, A_{pq}, ... nach 0 konvergieren). Ordnen wir daher jetzt zwei dyadische Mengen X, Y (desselben

Raumes oder verschiedener Räume) einander in der Weise zu, daß die Punkte x, y mit derselben Ziffernfolge $(p, q, r, \ldots)$ einander entsprechen, so hat diese Funktionsbeziehung die Eigenschaft, daß aus $\lim \overline{xx_n} = 0$ stets $\lim \overline{yy_n} = 0$ folgt und umgekehrt; in der üblichen Ausdrucksweise (Kap. IX) lassen sich also zwei dyadische Mengen umkehrbar eindeutig und beiderseits stetig aufeinander abbilden.

Betrachten wir jetzt eine allgemeine Menge der gleichen Art, wo nur statt der fortgesetzten Zweiteilung eine beliebige, auch beliebig wechselnde Anzahl von Teilmengen in Betracht gezogen wird. Also: es gebe π paarweise fremde Mengen A_p $(p = 1, 2, \ldots, \pi)$; für festes p gebe es $\varkappa_p$ paarweise fremde Mengen A_{pq} $(q = 1, 2, \ldots, \varkappa_p)$; für feste p, q gebe es ϱ_{pq} paarweise fremde Mengen A_{pqr} $(r = 1, 2, \ldots, \varrho_{pq})$ usw. Diese Zahlen seien $\geqq 2$; alle Mengen seien beschränkt, abgeschlossen, von Null verschieden; für jede der hiermit definierten Ziffernfolgen $(p, q, r, \ldots)$ sei wieder $A_p > A_{pq} > A_{pqr} > \ldots$ mit nach Null konvergenten Breiten, und X sei wie oben definiert. Wir wollen zeigen, daß dieses X dennoch nichts anderes als eine dyadische Menge ist, indem wir die A in gewisser Weise zusammenfassen und wieder spalten. Zur Abkürzung bedeute ξ einen endlichen Komplex von Ziffern $pqr\ldots$ (also p, pq, pqr, usw.), wobei wir uns die Komplexe gleicher Ziffernzahl lexikographisch geordnet denken, und η einen endlichen Komplex, der aus den Ziffern 1, 2 gebildet ist. Wir definieren dann die Mengen $B_1, B_2, B_{11}, \ldots$ durch Induktion folgendermaßen:

(α) $$B_1 = A_1 + A_3 + \cdots, \quad B_2 = A_2 + A_4 + \cdots.$$

(β) wenn B_η bereits als Summe von mindestens zwei Mengen A_ξ mit gleicher Ziffernzahl definiert und in lexikographischer Ordnung

$$B_\eta = A_\xi^{\mathrm{I}} + A_\xi^{\mathrm{II}} + A_\xi^{\mathrm{III}} + A_\xi^{\mathrm{IV}} + \cdots$$

ist, so sei

$$B_{\eta 1} = A_\xi^{\mathrm{I}} + A_\xi^{\mathrm{III}} + \cdots, \quad B_{\eta 2} = A_\xi^{\mathrm{II}} + A_\xi^{\mathrm{IV}} + \cdots.$$

(γ) wenn B_η einem einzigen A_ξ gleich ist, so sei

$$B_{\eta 1} = A_{\xi 1} + A_{\xi 3} + \cdots, \quad B_{\eta 2} = A_{\xi 2} + A_{\xi 4} + \cdots.$$

Wenn beispielsweise die folgenden Mengen A gegeben sind:

$$A_1 \; A_2 \; A_3;$$

$$A_{11} \; A_{12} \; A_{13}, \quad A_{21} \; A_{22} \; A_{23} \; A_{24}, \quad A_{31} \; A_{32}; \quad \text{usw.}$$

so ist zu setzen:

$$\begin{array}{llll} B_1 = A_1 + A_3 & , & B_2 = A_2; & \\ B_{11} = A_1, & B_{12} = A_3, & B_{21} = A_{21} + A_{23}, & B_{22} = A_{22} + A_{24}; \\ B_{111} = A_{11} + A_{13}, & B_{112} = A_{12}, & B_{121} = A_{31}, & B_{122} = A_{32}, \\ B_{211} = A_{21}, & B_{212} = A_{23}, & B_{221} = A_{22}, & B_{222} = A_{24}; \quad \text{usw.} \end{array}$$

Aus dieser Bestimmung folgt unmittelbar: jedes B_η ist entweder ein A_ξ oder eine Summe von solchen mit gleicher Ziffernzahl; es ist beschränkt, abgeschlossen und von Null verschieden; $B_{\eta 1}$ ist fremd zu $B_{\eta 2}$ (auch B_1 zu B_2); $B_\eta \geqq B_{\eta 1} + B_{\eta 2}$, also $B_\eta > B_{\eta 1}$ und $B_\eta > B_{\eta 2}$. Jede Folge $B_i, B_{ik}, B_{ikl}, \ldots$ enthält eine Teilfolge $A_p, A_{pq}, A_{pqr}, \ldots$; die Breiten der B nehmen nicht zu, konvergieren also mit den Breiten der A nach 0, und die B erzeugen eine dyadische Menge Y; da der Durchschnitt einer absteigenden Folge gleich dem einer Teilfolge ist, so ist jeder Punkt y zugleich ein Punkt x. Umgekehrt ist jedes A_ξ ein B_η, jeder Punkt x also Durchschnittspunkt einer absteigenden Folge von Mengen B_η, diese wieder Teilfolge einer Folge $B_i, B_{ik}, B_{ikl}, \ldots$, und demnach jeder Punkt x auch ein Punkt y. Daher ist $X = Y$, die Menge X ist dyadisch.

Jede kompakte perfekte punkthafte Menge eines metrischen Raumes ist dyadisch. Sie zerfällt nämlich für hinlänglich kleines ϱ in mindestens zwei ϱ-Komponenten (S. 301) $A = A_1 + A_2 + \ldots + A_\pi = \Sigma A_p$; für hinlänglich kleines $\varrho_1 < \varrho$ zerfällt sie in der Weise in ϱ_1-Komponenten, daß auch jedes einzelne $A_p = \sum\limits_q A_{pq}$ deren mindestens zwei enthält, usf. Ferner folgt aus § 6, VII, daß für eine absteigende Folge kompakter abgeschlossener nichtverschwindender Mengen deren Breite nach der Breite ihres Durchschnitts konvergiert; in unserem Fall konvergieren also die Breiten von $A_p, A_{pq}, A_{pqr}, \ldots$ nach 0, da ihr Durchschnitt eine Komponente von A ist, d. h. aus einem einzigen Punkte besteht. Die Menge

$$A = \mathfrak{D}(\Sigma A_p, \ \Sigma A_{pq}, \ \ldots)$$

ist also dyadisch. Zwei solche Mengen sind daher immer eindeutige stetige Bilder voneinander.

Die Komponentenmenge $\mathfrak{A}$ einer abgeschlossenen kompakten Menge A ist punkthaft; denn sind P, Q zwei verschiedene Komponenten, so gibt es, weil die Komponenten auch Quasikomponenten sind, eine Zerlegung in zwei abgeschlossene Mengen

$$A = X + Y, \quad X = P + P' + \ldots, \quad Y = Q + Q' + \ldots,$$

und dieser entspricht, im Sinne der Distanzen, eine Zerlegung

$$\mathfrak{A} = \mathfrak{X} + \mathfrak{Y}, \quad \mathfrak{X} = \{P, P', \ldots\}, \quad \mathfrak{Y} = \{Q, Q', \ldots\}$$

in zwei ebenfalls abgeschlossene Komponentenmengen, da jedes Element von $\mathfrak{X}$ von jedem Element von $\mathfrak{Y}$ eine untere Entfernung $\geqq \delta(X, Y)$ hat. P, Q gehören also keiner zusammenhängenden Teilmenge $\{P, Q, \ldots\}$ von $\mathfrak{A}$ an.

Zwei abgeschlossene kompakte Mengen mit perfekten Komponentenmengen, d. h. ohne isolierte Komponenten, lassen sich also

komponentenweise eineindeutig und beiderseits stetig aufeinander beziehen, d. h. jeder Komponente P der einen entspricht eine Komponente Q der andern so, daß aus $\lim \delta(P, P_n) = 0$ auch $\lim \delta(Q, Q_n) = 0$ folgt und umgekehrt.

Auf der geraden Linie enthält jede zusammenhängende Menge, wenn sie nicht aus einem Punkte besteht, eine Strecke, also innere Punkte; eine Randmenge ist hier stets punkthaft. (Daß umgekehrt eine punkthafte Menge Randmenge ist, gilt in jedem euklidischen Raum, da ein Kugelinneres zusammenhängend ist, § 10.) Unsere Beispiele von perfekten nirgendsdichten Mengen (S. 253) geben also auch solche von punkthaften Mengen; damals hatten wir allerdings auf die Beschränktheit keinen Wert gelegt, die man durch Schnitt mit einem geeigneten Intervall leicht herstellen kann. Um sie als dyadische Mengen zu konstruieren, kann man unter den $A_p, A_{pq}, \ldots$ einfach abgeschlossene Intervalle (a, b) (d. h. die Menge der Zahlen x mit $a \leqq x \leqq b$, $a < b$) verstehen. Wählt man z. B. $A_{\xi 1}$, $A_{\xi 2}$ als das linke und rechte Drittel von A_ξ, und beginnt mit

$$A_1 = (0, \tfrac{1}{3}), \quad A_2 = (\tfrac{2}{3}, 1),$$

also weiter

$$A_{11} = (0, \tfrac{1}{9}),\ A_{12} = (\tfrac{2}{9}, \tfrac{1}{3}),\ A_{21} = (\tfrac{2}{3}, \tfrac{7}{9}),\ A_{22} = (\tfrac{8}{9}, 1) \quad \text{usf.},$$

so wird X die Cantorsche Menge der Zahlen

$$x = \frac{x_1}{3} + \frac{x_2}{3^2} + \frac{x_3}{3^3} + \cdots \qquad (x_n = 0, 2)$$

in deren triadischer Entwicklung keine Eins vorkommt. Bezeichnet man die Länge des Intervalls A_ξ mit δ_ξ und den Limes der abnehmenden Zahlen $\Sigma\delta_p$, $\Sigma\delta_{pq}$, ... als das „Längenmaß" der Menge X (vgl. Kap. X), so ist dieses bei der Cantorschen Menge 0, nämlich der Limes von $\frac{2}{3}, (\frac{2}{3})^2, (\frac{2}{3})^3, \ldots$; es ist aber evident und wurde schon damals von uns bemerkt, daß man auch positives Längenmaß erhalten kann, indem man die Intervalllängen langsam genug abnehmen läßt, z. B. mit einer beliebigen Folge abnehmender, nach $\varrho > 0$ konvergierender positiver Zahlen ϱ_n die Intervalle $A_p, A_{pq}, A_{pqr}, \ldots$ von den Längen $\frac{1}{2}\varrho_1, \frac{1}{4}\varrho_2, \frac{1}{8}\varrho_3, \ldots$ wählt.

Das Gleiche gilt in der Ebene; zeichnet man z. B. (Fig. 12) in die Ecken eines Quadrats[1] A vier kleinere Quadrate A_1, A_2, A_3, A_4, die das große Quadrat nicht ausfüllen, sondern ein Kreuz freilassen, in jedes A_p wieder vier solche Quadrate A_{pq}, in jedes A_{pq} abermals vier Quadrate A_{pqr} usw., so bekommt die punkthafte Menge X ein positives Flächenmaß bei genügend langsamer Abnahme der Quadratseiten oder bei genügender Schmalheit der freibleibenden Kreuze,

[1] Darunter ist die Quadratfläche, Inneres und Umfang, zu verstehen.

z. B. indem man die Quadratseiten wie oben $= \frac{1}{2}\varrho_1, \frac{1}{4}\varrho_2, \frac{1}{8}\varrho_3, \ldots$ wählt. Dies ist noch erstaunlicher als daß eine nirgendsdichte Menge (die ja in der Ebene nicht punkthaft zu sein braucht) positives Flächenmaß haben kann.

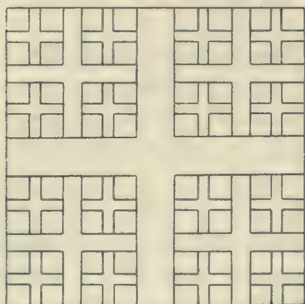


Fig. 12.

In einem vollständigen Raum können wir ferner noch die Ergebnisse von Kap. VII, § 8 über relative Dichtigkeit vervollständigen. Wir sahen damals: wenn A_1, A_2 Relativgebiete von M und in M dicht sind, so ist auch ihr Durchschnitt noch in M dicht; und wenn die beiden beliebigen Mengen A_1, A_2 zu M nirgendsdicht sind, so ist auch ihre Summe noch zu M nirgendsdicht. Dies läßt sich jetzt in nachstehender Weise auf Mengenfolgen übertragen.

V. In einem vollständigen Raume sei $A_1, A_2, \ldots$ eine Folge von Mengen, deren jede zur vorangehenden dicht ist. Ist dann G_n ein Gebiet $\geqq A_n$, so ist auch noch $G_\delta = \mathfrak{D}(G_1, G_2, \ldots)$ zu A_1 dicht.

Wir umgeben einen Punkt a_1 von A_1 mit einer abgeschlossenen Kugel V_1. Im entsprechenden U_1 liegt sicher ein Punkt a_2 von A_2; wir umgeben ihn mit einer abgeschlossenen Kugel $V_2 \leqq V_1$. Im entsprechenden U_2 liegt sicher ein Punkt a_3 von A_3, den wir mit einer abgeschlossenen Kugel $V_3 \leqq V_2$ umgeben usw. Lassen wir die Radien dieser Kugeln nach Null konvergieren und wählen gleichzeitig $V_1 \leqq G_1, V_2 \leqq G_2, \ldots$, so reduziert sich $\mathfrak{D}(V_1, V_2, \ldots)$ auf einen zu G_δ gehörigen Punkt x. Da V_1 beliebig klein gewählt werden kann, so liegt in jeder Umgebung von a_1 ein Punkt x von G_δ; G_δ ist zu A_1 dicht.

VI. In einem vollständigen Raume sei $A_1, A_2, \ldots$ eine Folge paarweise kongruenter (zueinander dichter) Mengen G_δ; dann ist auch ihr Durchschnitt $A = \mathfrak{D}(A_1, A_2, \ldots)$ noch mit ihnen kongruent.

Es sei $A_p = \mathfrak{D}(G_{p1}, G_{p2}, \ldots)$; wir verwandeln die Doppelfolge der Gebiete G_{pq} nach dem bekannten Diagonalverfahren in eine einfache Folge

$$G_1 = G_{11},\ G_2 = G_{12},\ G_3 = G_{21}, \ldots,\ G_n = G_{pq}, \ldots$$

und setzen $B_n = A_p$, so daß $G_n \geqq B_n$ und $A = \mathfrak{D}(G_1, G_2, \ldots)$ wird. Die paarweise kongruenten B_n erfüllen reichlich die Voraussetzung des vorigen Satzes über die A_n; also ist A zu $B_1 = A_1$ dicht oder in A_1 dicht oder mit A_1 kongruent.

Wir können auch sagen: sind die A_n mit einer und derselben Menge M kongruent, so ist auch A noch mit M kongruent, und insbesondere: sind die A_n in einer Menge M dicht, so ist auch A noch in M dicht. Wir bemerken davon den Spezialfall:

VII. Ist M eine Menge G_δ in einem vollständigen Raume und bilden $A_1, A_2, \ldots$ eine Folge in M dichter Relativgebiete, so ist auch ihr Durchschnitt $A = \mathfrak{D}(A_1, A_2, \ldots)$ noch in M dicht.

Denn die Mengen A_n sind dann selbst von der Form G_δ. Daß der Satz ohne die Voraussetzung $M = G_\delta$ nicht zu gelten braucht, zeigt das Beispiel der Menge $M = \{r_1, r_2, \ldots\}$ der rationalen Punkte eines euklidischen Raumes: hier ist $A_n = M - \{r_n\}$ in M dicht und Relativgebiet, nämlich Komplement einer endlichen, also abgeschlossenen Menge, und dennoch $\mathfrak{D}(A_1, A_2, \ldots) = 0$. M ist in diesem Fall eben keine Menge G_δ.

Der in VII genannte Durchschnitt A ist ein in M dichtes G_δ. Nun kann man für ein in M dichtes A auf das Nichtverschwinden des Kerns A_k jedenfalls dann schließen, wenn die Menge M_j der isolierten Punkte von M nicht in M dicht ist (denn wenn A separiert und in M dicht ist, so ist $M_j = A_j$ in A, also auch in M dicht); z. B. wenn M insichdicht und von Null verschieden ist. Unter dieser Einschränkung ist A also immer noch mindestens von der Mächtigkeit des Kontinuums.

Aus VII folgt nun:

VIII. Ist M eine Menge G_δ in einem vollständigen Raume und bilden $A_1, A_2, \ldots$ eine Folge in M nirgendsdichter Mengen, so ist das Komplement $M - A$ der Summe $A = \mathfrak{S}(A_1, A_2, \ldots)$ noch in M dicht.

Erinnern wir uns aus Kap. VII, § 8 der zu zwei Mengen A, M gehörigen Mengen

$$P = \mathfrak{D}(M, A_\alpha), \quad Q = M - P;$$

daß A zu M nirgendsdicht ist, ist gleichbedeutend damit, daß Q in M dicht ist. P ist in M abgeschlossen, Q ein Relativgebiet in M. Ist $A \leqq M$, so ist $A \leqq P$. Bezeichnen wir jetzt, unter den Voraussetzungen von VIII, die der Menge A_n entsprechenden Mengen mit P_n, Q_n und setzen

$$P = \mathfrak{S}(P_1, P_2, \ldots), \quad Q = \mathfrak{D}(Q_1, Q_2, \ldots) = M - P,$$

so sind die Q_n in M dichte Relativgebiete, also auch Q noch in M dicht; überdies ist $A_n \leqq P_n$, $A \leqq P$, $M - A \geqq Q$, also auch $M - A$ in M dicht.

Der Leser beachte wohl, daß wir nicht behauptet haben und nicht behaupten können, die Summe A sei (wie im Fall endlicher Zahl der Summanden) noch in M nirgendsdicht; sie kann sogar

in M dicht sein, aber jedenfalls ist ihr Komplement $M - A$ in M dicht. Ist z. B. $A = \{r_1, r_2, \ldots\}$ die Menge der rationalen Punkte eines euklidischen Raumes E, so ist $A_n = \{r_n\}$ in E nirgendsdicht, A in E dicht, $E - A$ aber jedenfalls auch.

Ist in VIII M nicht Null, so ist A nie mit M identisch; ist M_j nicht in M dicht, so ist, wie aus der entsprechenden Bemerkung zu VII folgt, $M - A$ noch mindestens von der Mächtigkeit des Kontinuums.

Man nennt (nach R. Baire) die Summe einer Folge von Mengen, die in M nirgendsdicht sind, eine Menge erster Kategorie in bezug auf M, und eine Teilmenge von M, die nicht von erster Kategorie ist, eine Menge zweiter Kategorie. Schreiben wir dafür kurz M_1, M_2. Jede Teilmenge eines M_1 ist wieder ein M_1; die Summe endlich oder abzählbar vieler M_1 ist wieder ein M_1. Wenn die ganze Menge M in bezug auf sich selbst von erster Kategorie ist, so sind alle ihre Teilmengen Mengen M_1 und es gibt kein M_2. Ist umgekehrt M ein M_2, so ist auch $M - M_1$ stets ein M_2. In dieser Bezeichnung sagt der Satz VIII, daß in einem vollständigen Raume eine Menge M, die ein G_δ und nicht Null ist, stets ein M_2, und daß $M - M_1$ stets ein M_2 und in M dicht ist.

Wenn $A = \mathfrak{S}(A_1, A_2, \ldots) \leqq M$ und $M - A$ in M dicht ist, wenn überdies die Mengen A_n in M abgeschlossen sind, so ist A ein M_1. Denn die oben definierten Mengen P_n, Q_n sind dann resp. $= A_n$, $M - A_n$, und da $M - A_n \geqq M - A$ in M dicht ist, so ist A_n in M nirgendsdicht. Wir können dies so ausdrücken: ist A der Durchschnitt von M mit einer Menge F_σ und $M - A$ in M dicht, so ist A in bezug auf M von erster Kategorie.[1]

Eine Summe A von abzählbar vielen abgeschlossenen Randmengen A_n (die ja nach S. 253 nirgendsdicht sind) ist in bezug auf den Raum E von erster Kategorie; ist E vollständig, so ist $E - A$ dicht, also auch A noch eine Randmenge. Ohne die Abgeschlossenheit der A_n gilt dies keineswegs; schon die Summe zweier Randmengen kann innere Punkte haben (S. 217).

§ 10. Euklidische Räume.

In einem euklidischen Raume $E = E_n$, d. h. in der Menge der reellen Zahlenkomplexe $x = (x_1, x_2, \ldots, x_n)$ mit den Entfernungen

$$\overline{xy} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} \geqq 0$$

[1] Wir werden sehen, daß die Menge der Unstetigkeitspunkte einer in M definierten Funktion stets eine Menge $A = \mathfrak{D}(M, F_\sigma)$ ist; ist die Funktion nur „punktweise“ unstetig, d. h. liegt die Menge $M - A$ ihrer Stetigkeitspunkte in M dicht, so ist A in bezug auf M von erster Kategorie.

gelten sämtliche bisher abgeleiteten Resultate. Denn E ist ein metrischer Raum, seine sphärischen Umgebungen erfüllen also die Axiome (A) bis (D) und, bei Beschränkung auf rationale Radien, das erste Abzählbarkeitsaxiom (E). In E ist eine abzählbare Menge dicht, also gilt, bei Ersetzung der sphärischen Umgebungen durch gleichwertige, das zweite Abzählbarkeitsaxiom (F). Endlich ist E vollständig, und zwar sind hier kompakte, total beschränkte und beschränkte Mengen identisch (die Identität zwischen total beschränkten und schlechthin beschränkten Mengen beruht auf der Möglichkeit der Würfelteilung, S. 312), sodaß auch im n-dimensionalen Raume wie im eindimensionalen der Satz gilt:

I (Satz von Bolzano-Weierstrass). Jede beschränkte unendliche Menge eines euklidischen Raumes hat mindestens einen Häufungspunkt.

Für die Struktur eines Raumes sind vor allem seine zusammenhängenden Teilmengen charakteristisch. Da wir zwar schon viele Sätze über Zusammenhang aufgestellt haben, aber tatsächlich zusammenhängende Mengen, die nicht nur aus einem Punkte bestehen, bisher nur auf der geraden Linie (Kap. VII, § 9) kennen, so ist begreiflich, daß auch im euklidischen Raume die Geraden und Strecken eine bedeutende Rolle spielen werden. Was wir darunter verstehen, ist ja evident: sind $a = (a_1, a_2, \ldots, a_n)$ und $b = (b_1, b_2, \ldots, b_n)$ zwei verschiedene Punkte, so ist die durch sie bestimmte Gerade die Menge der Punkte $x = (x_1, x_2, \ldots, x_n)$, die sich mit reellem t in der Form

$$x_i = t a_i + (1 - t) b_i \qquad (i = 1, 2, \ldots, n)$$

darstellen lassen, und die durch sie bestimmte (abgeschlossene) Strecke, die wir kurz (a, b) nennen wollen, die Menge der Punkte jener Geraden, für die $0 \leqq t \leqq 1$. Für $b = a$ soll die Strecke (a, a) aus dem einzigen Punkt a bestehen. Jede Strecke (a, b) ist zusammenhängend, denn sie ist es jedenfalls als Teilmenge ihrer eigenen Geraden, und nach den Relativitätsbetrachtungen (Kap. VII, § 6) ist sie es auch als Teilmenge von E.[1]

Jede konvexe Menge, d. h. eine solche, die mit zwei verschiedenen Punkten a, b auch alle Punkte der Strecke (a, b) enthält, ist hiernach zusammenhängend, so vor allem der ganze Raum E selbst. Danach ist eine abgeschlossene Menge niemals ein Gebiet (außer E und der Nullmenge), da sonst ihr Komplement abgeschlossen und E in zwei abgeschlossene Mengen zerlegbar wäre. Sodann ist

---

[1] Man sieht leicht, wie man auch im Hilbertschen Raume und in den Funktionenräumen (§ 5) gerade Linien und Strecken definieren und damit die Betrachtungen des Textes auf sie übertragen kann.

jede sphärische Umgebung (das Innere einer Kugel) zusammenhängend, und diese Bemerkung, so einfach sie ist, hat weitreichende Konsequenzen. Zunächst folgt:

II. Im euklidischen Raum sind die Komponenten eines Gebietes selbst Gebiete; jedes nichtverschwindende Gebiet ist die Summe von endlich oder abzählbar vielen, paarweise fremden, zusammenhängenden Gebieten.

Denn ist x ein Punkt des Gebietes G, U_x eine Umgebung $\leqq G$, so ist U_x selbst zusammenhängend, also Teilmenge der zu x gehörigen Komponente, demnach innerer Punkt dieser Komponente. Das Weitere folgt aus § 3, III.

Durch diesen Satz ist im eindimensionalen Raum $E_1 = T$ die Struktur der Gebiete G und damit auch der abgeschlossenen Mengen F vollkommen geklärt. G ist, wenn es nicht 0 oder T ist, eine Summe von endlich oder abzählbar vielen, paarweise fremden offenen Strecken, wozu eventuell ein oder zwei offene Halbgerade treten können. $F = T - G$ ist die allgemeinste abgeschlossene Menge $(\neq 0, T)$; sie wird perfekt dann und nur dann, wenn keine zwei Strecken von G mit einem Endpunkt aneinander stoßen (der ein isolierter Punkt von F sein würde). Beispiele sind uns aus Kap. VII, § 8 bekannt.

Wenn wir im euklidischen Raume irgend eine Zerlegung

$$G = G_1 + G_2 + \cdots$$

des Gebietes G in paarweise fremde, zusammenhängende Gebiete G_n kennen, so sind diese die Komponenten von G. Denn eine zusammenhängende Menge $C \leqq G$ kann nicht mit zweien dieser Gebiete, etwa G_1 und G_2, Punkte gemein haben, da $\Gamma_1 = G - G_1$ selbst ein Gebiet ist und $C = \mathfrak{D}(C, G_1) + \mathfrak{D}(C, \Gamma_1)$ eine Zerlegung in zwei Relativgebiete wäre. Ebenso zeigt die Zerlegung $G = G_1 + \Gamma_1$ in zwei Gebiete oder zwei relativ abgeschlossene Mengen, daß G_1 und G_2 verschiedenen Quasikomponenten (S. 248) von G angehören, daß auch bei einem Gebiet also, wie bei abgeschlossenen kompakten (beschränkten) Mengen, Komponenten und Quasikomponenten identisch sind.

III. Ist H die Grenze des zusammenhängenden Gebiets G, so ist G eine Komponente von $E - H$.

Denn $G + H$ ist abgeschlossen, $E - (G + H)$ ein Gebiet; ist dessen Komponentenzerlegung $G_1 + G_2 + \ldots$, so ist $E - H = G + G_1 + G_2 + \ldots$, G Komponente von $E - H$. Ein zusammenhängendes Gebiet ist also durch Angabe eines seiner Punkte sowie seiner Grenze H bestimmt, als die den genannten Punkt enthaltende Komponente von $E - H$.

Der Zusammenhang des Kugelinneren erlaubt auch, einige frühere Formeln (Kap. VII, § 2, (5) und (11)) über Rand und Grenze einer Summe von Mengen zu vervollständigen. Es gilt:

IV. Der Rand einer Menge ist die Summe der Ränder ihrer Komponenten.

In der Tat sei

$$A = P + Q + \dots$$

die Zerlegung von A in (endlich oder unendlich viele) Komponenten. Wie wir damals sahen, ist

$$A_r \leqq P_r + Q_r + \dots .$$

Andererseits sei p ein Randpunkt von P; wir behaupten, daß er auch Randpunkt von A ist. Denn wäre er innerer Punkt von A und $U_p \leqq A$, so liegt die zusammenhängende Menge U_p in einer Komponente von A, also in P, und p wäre innerer Punkt von P. Demnach ist $P_r \leqq A_r$ und

$$A_r \geqq P_r + Q_r + \cdots ,$$

also

$$A_r = P_r + Q_r + \dots .$$

Das ist die behauptete Formel, aus der auch für die inneren Punkte

$$A_i = P_i + Q_i + \dots$$

folgt.

Ist B das Komplement von A (wir setzen A von 0 und E verschieden voraus) und P eine Komponente von A, so ist nach den letzten Formeln wegen $B_\alpha = B + A_r$

$$\mathfrak{D}(P, B_\alpha) \doteq \mathfrak{D}(P, A_r) = P_r = P - P_i$$

und, weil P in A abgeschlossen ist,

$$\mathfrak{D}(B, P_\alpha) = \mathfrak{D}(E - A, P_\alpha) = P_\alpha - P.$$

Diese beiden Mengen können nicht gleichzeitig verschwinden, da sonst P gleichzeitig abgeschlossen und ein Gebiet wäre, was mit $0 < P < E$ unvereinbar ist. Also muß wenigstens eine der beiden Mengen B, P Häufungspunkte der andern enthalten. Ist B insbesondere (wie P) zusammenhängend, so ist daher auch $B + P$ zusammenhängend; und bezeichnet man mit $P, P', P'', \dots$ irgendwelche, endlich oder unendlich viele Komponenten von A, so ist auch

$$\mathfrak{S}(B + P, B + P', B + P'', \dots) = B + P + P' + P'' + \dots$$

zusammenhängend, als Summe zusammenhängender Mengen mit nichtverschwindendem Durchschnitt. Also:

V. Eine zusammenhängende Menge bleibt zusammenhängend, wenn man ihr beliebig viele Komponenten ihres Komplements hinzufügt.

Ist ferner

$$G = G_1 + G_2 + \dots$$

eine Summe paarweise fremder Gebiete (die nicht notwendig die Komponenten von G zu sein brauchen) und bezeichnet man die Komplemente dieser Gebiete mit F, F_n, die Grenzen mit H, H_n, so gilt statt der früheren Ungleichung

$$H \geqq S = \mathfrak{S}(H_1, H_2, \dots)$$

jetzt die genaue Formel

$$H = S_\alpha.$$

Denn sei x ein Punkt von H; eine Umgebung U_x enthält also einen Punkt von G, etwa von G_1, und da sie auch einen nicht zu G_1 gehörigen Punkt (x) enthält und zusammenhängend ist, so muß sie auch einen Punkt der Grenze H_1, also von S enthalten. Demnach ist x ein α-Punkt von S, $H \leqq S_\alpha$, und aus $S \leqq H \leqq S_\alpha$ folgt durch Bildung der α-Menge, weil H abgeschlossen ist, $H = S_\alpha$. Also:

VI. Die Grenze einer Summe paarweise fremder Gebiete besteht aus den Grenzpunkten der einzelnen Gebiete und den Häufungspunkten dieser Grenzpunkte.

Für eine endliche Gebietssumme $G = G_1 + G_2 + \dots + G_n$ ist $H = \mathfrak{S}(H_1, H_2, \dots, H_n)$.

Z. B. ist im linearen Raum jedes Gebiet G die Summe von offenen Strecken G_n (worunter sich auch ein oder zwei Halbgerade befinden können), H_n ist das Paar der Endpunkte dieser Strecke (resp. der eine Endpunkt der Halbgeraden), und die Grenze von G ist die Menge dieser Endpunkte und ihrer Häufungspunkte.

Unter einem Streckenzug $(a_1, a_2, \dots, a_m)$ verstehen wir die Summe der Strecken $(a_1, a_2), (a_2, a_3), \dots, (a_{m-1}, a_m)$; auch eine solche Menge ist zusammenhängend, denn nach Kap. VII, § 7, II sind der Reihe nach (a_1, a_2, a_3), (a_1, a_2, a_3, a_4) usw. zusammenhängend. Eine Menge, in der zwei Punkte durch einen zur Menge gehörigen Streckenzug verbunden werden können, ist zusammenhängend. Davon gilt bei Gebieten auch die Umkehrung:

VII. In einem zusammenhängenden Gebiet G können je zwei Punkte durch einen zu G gehörigen Streckenzug verbunden werden.[1]

Ist a_0 ein Punkt von G, so unterscheiden wir in G diejenigen Punkte a, die sich mit a_0 durch einen zu G gehörigen Streckenzug $(a_0, \dots, a)$ verbinden lassen, von denjenigen Punkten b, die das nicht

[1] Bei Weierstrass ist diese Eigenschaft mit in die Definition des Gebietes aufgenommen.

tun. Die bezüglichen Mengen seien A und B, also $G = A + B$. Beide Mengen sind Gebiete. Denn ist x ein Punkt von $U_a \leqq G$, so ist $(a_0, \ldots, a, x)$ ein Streckenzug, der a_0 mit x verbindet, also gehört U_a zu A, A ist ein Gebiet. Umgekehrt, ist y ein Punkt von $U_b \leqq G$, und wäre y mit a_0 durch $(a_0, \ldots, y)$ verbunden, so ließe sich auch b mit a_0 durch $(a_0, \ldots, y, b)$ verbinden; da dies nicht sein darf, so ist y ein Punkt von B, $U_b \leqq B$, B ein Gebiet. Da nun $A > 0$ (weil es den Punkt a_0 enthält), so muß für ein zusammenhängendes Gebiet $B = 0$ sein, d. h. alle Punkte sind mit a_0 und demnach auch untereinander verbindbar.

Wir geben einige Mengen an, deren Zusammenhang durch Streckenzüge bewiesen werden kann. In der Ebene (oder in einem mehr als zweidimensionalen Raume, aber nicht auf der geraden Linie) ist die Menge der irrationalen Punkte oder allgemeiner jede durch Tilgung einer höchstens abzählbaren Menge R entstehende Menge $E - R$ zusammenhängend. Denn sind a, b zwei Punkte von $E - R$, so gibt es unter den $\aleph$ Geraden durch a höchstens abzählbar viele, die einen Punkt von R treffen, also immer noch $\aleph$ solche, die ganz in $E - R$ verlaufen; zwei solche, einander schneidende Gerade durch a, b geben einen diese Punkte verbindenden Streckenzug (a, x, b). Auch ein Kreisinneres bleibt nach Wegnahme einer höchstens abzählbaren Menge zusammenhängend. Der dreidimensionale Raum bleibt nach Tilgung von endlich oder abzählbar vielen Geraden zusammenhängend; denn sind a, b zwei Punkte der übrigbleibenden Menge, so gibt es unter den $\aleph$ Ebenen durch a, b höchstens abzählbar viele, die eine der Geraden ganz enthalten, also immer noch $\aleph$ solche, die jede Gerade in höchstens einem Punkte treffen, und in einer dieser Ebenen lassen sich a, b durch einen Streckenzug verbinden.

Eine zusammenhängende Menge eines mindestens zweidimensionalen Raumes bleibt nach Tilgung eines inneren Punktes zusammenhängend. Denn wäre $A - \{a\} = P + Q$ in zwei relativ abgeschlossene Mengen zerlegbar, so gibt es eine zu A gehörige Umgebung U_a; $U_a - \{a\}$ ist noch zusammenhängend, gehört also einem der Summanden an, etwa zu P; dann wären aber auch $P + \{a\}$ und Q in A abgeschlossen, da a kein Häufungspunkt von Q ist. Auch die Wegnahme zweier oder endlich vieler[1] innerer Punkte $a, b, \ldots$ hebt den Zusammenhang von A nicht auf, da ja b noch innerer Punkt von $A - \{a\}$ ist usw.

---

[1] Sogar abzählbar vieler, wie wir ohne Beweis bemerken; überhaupt weist eine eingehendere Betrachtung noch viel allgemeinere Mengen nach, durch deren Tilgung eine zusammenhängende Menge nicht zerfällt.

Wir sagen, die beiden verschiedenen Punkte x, y werden durch C verbunden, wenn C zusammenhängend ist und x, y enthält. Von einer die Punkte x, y selbst nicht enthaltenden Menge D sagen wir, daß sie x und y trennt, wenn jede x und y verbindende Menge die Menge D trifft (d. h. mit ihr mindestens einen Punkt gemein hat); also nicht trennt, wenn es eine die Punkte x, y verbindende und die Menge D nicht treffende Menge gibt. Die Punkte x, y werden durch D getrennt oder nicht, je nachdem sie zwei verschiedenen Komponenten oder derselben Komponente von $E - D$ angehören.

Z. B. werden in der geraden Linie zwei Punkte schon durch eine Menge aus einem einzigen (zwischenliegenden) Punkt getrennt, in der Ebene erst durch eine Gerade, ein geschlossenes Polygon, einen Kreis u. dgl. (Verstehen wir unter E eine Kreisperipherie, so ist zur Trennung eine Menge aus mindestens zwei Punkten notwendig und hinreichend.)

Wenn eine abgeschlossene Menge F die Punkte x, y nicht trennt, so gehören x, y einer Komponente von $E - F$, also einem zusammenhängenden Gebiet $\leqq E - F$ an und lassen sich nach VII durch einen F nicht treffenden Streckenzug S verbinden. Aus dem Umstand, daß S eine abgeschlossene beschränkte Menge ist, folgt:

VIII. Wenn die abgeschlossenen Mengen $F_1 \geqq F_2 \geqq \ldots$ die Punkte x, y trennen, so tut es auch ihr Durchschnitt $F = \mathfrak{D}(F_1, F_2, \ldots)$.

Denn andernfalls ließen sich x, y durch einen Streckenzug S verbinden, der F nicht trifft, während S als zusammenhängende Menge jedes F_n treffen muß. Die abgeschlossenen beschränkten Mengen $\mathfrak{D}(F_n, S)$ sind also von Null verschieden und bilden eine absteigende Folge; dann müßte aber nach dem Cantorschen Satze (S. 230) auch ihr Durchschnitt $\mathfrak{D}(F, S)$ von Null verschieden sein, d. h. S doch die Menge F treffen.

Offenbar gilt dies auch von jeder absteigend wohlgeordneten Menge von abgeschlossenen Mengen (§ 4).

Sind A, B zwei fremde, nichtverschwindende Mengen, so werden je zwei innere Punkte von A und B durch die Grenze A_g oder durch die Grenze B_g getrennt (Kap. VII, § 7, VII). Sind A, B speziell Gebiete, so werden je zwei Punkte von A und B durch eine dieser Grenzen getrennt (oder auch durch das Komplement $E - (A + B)$, oder durch dessen Rand, d. h. durch die Grenze des Gebietes $A + B$).

Zu zwei fremden, nichtverschwindenden Mengen A, B, die in $A + B$ abgeschlossen sind, von denen also keine einen Häufungspunkt der andern enthält, gibt es ebenfalls eine abgeschlossene

Menge, die je zwei Punkte von A und B trennt. Man kann nämlich A und B in zwei fremde Gebiete $G \geqq A$, $H \geqq B$ einschließen, wonach die Grenze des einen oder das Komplement $E-(G+H)$ das Verlangte leistet. In der Tat: die untere Entfernung eines Punktes a von B ist stets positiv; setzen wir $\delta(a, B) = 2\varrho_a$, ebenso $\delta(b, A) = 2\varrho_b$, so haben die Umgebungen U_a, U_b mit den Radien ϱ_a, ϱ_b keinen Punkt gemein, weil

$$\overline{ab} \geqq 2\varrho_a, \quad \overline{ab} \geqq 2\varrho_b, \quad \overline{ab} \geqq \varrho_a + \varrho_b,$$

und die Gebiete $G = \mathfrak{S}_a U_a$, $H = \mathfrak{S}_b U_b$ sind fremd.

Ist eine der Mengen, etwa A, beschränkt, so kann man erreichen, daß auch die trennende Menge beschränkt wird, indem man vor der Bildung jener Gebiete zu B noch das Äußere einer Kugel hinzufügt, in deren Innerem A liegt.

Bei den weiteren Betrachtungen ziehen wir vor, uns auf die Ebene zu beschränken. Die Übertragung auf den drei- oder mehrdimensionalen Raum bietet zum Teil nicht unbeträchtliche Schwierigkeiten dar, weil die Rolle, die in der Ebene die Polygone spielen, dort den viel weniger einfachen Polyedern und Polytopen zufällt. Selbst in der Ebene werden wir finden, daß die anscheinend plausibelsten Aussagen der „Anschauung" ziemlich umständliche Beweise verlangen. Eine gewisse Weitläufigkeit wird schon dadurch herbeigeführt, daß nichtkompakte, hier also unbeschränkte Mengen sich in vielen Beziehungen anders und zwar weniger regelmäßig verhalten als beschränkte. Ein Radikalmittel dagegen wäre die Adjunktion eines „unendlich fernen" Punktes, wie in der Funktionentheorie: aber dadurch zerstört man den Charakter eines metrischen Raumes, und wenn man diesen Übelstand wieder durch stereographische Abbildung auf die Kugel beseitigt, so gehen dafür gewisse elementargeometrische Vorzüge der Ebene verloren. Wir müssen uns also mit den „dreary infinities of homaloidal space", wie Clifford sagt, so gut es geht, abzufinden suchen.

§ 11. Die euklidische Ebene.

Wir haben einen Streckenzug $S = (a_0, a_1, a_2, \ldots, a_m)$ als Summe der abgeschlossenen Strecken $(a_0, a_1), (a_1, a_2), \ldots, (a_{m-1}, a_m)$ erklärt. Von dem Fall, daß S sich auf einen einzigen Punkt reduziert, sehen wir ab und können die Endpunkte jeder Strecke als verschieden annehmen $(a_{i-1} \neq a_i)$. Eine Menge, die überhaupt ein Streckenzug ist, kann immer auf unendlich viele Weisen

$$S = (a_0, a_1, \ldots, a_m) = (b_0, b_1, \ldots, b_n) = \ldots$$

als solcher dargestellt werden; die kleinste der hierbei auftretenden Zahlen $m, n, \ldots$ heiße die Streckenzahl von S. Wenn S „sich selbst nicht kreuzt“ oder, präziser, wenn die Streckenzüge $(a_0, a_1, \ldots, a_i)$ und $(a_i, a_{i+1}, \ldots, a_m)$ für jedes $i = 1, 2, \ldots, m-1$ nur den Punkt a_i gemein haben, wozu wir auch den Fall einer einzigen Strecke $S = (a_0, a_1)$ rechnen, so heißt S ein einfacher Streckenzug oder ein Weg. Wird ein Weg von der Streckenzahl m in der Form $(a_0, a_1, \ldots, a_m) = (b_0, b_1, \ldots, b_m)$ dargestellt, so sind die Punkte b mit den Punkten a in derselben oder der umgekehrten Reihenfolge identisch; sie heißen die Eckpunkte, insbesondere a_0, a_m die Endpunkte des Weges und sind paarweise verschieden; drei konsekutive Eckpunkte a_{i-1}, a_i, a_{i+1} liegen nicht in gerader Linie. Sind x, y zwei verschiedene Punkte eines Streckenzuges S, und ist unter den Streckenzügen $S' \leqq S$, die x, y enthalten, W ein solcher mit kleinster Streckenzahl, so ist W ein Weg und es gibt auch jedenfalls einen solchen Weg mit den Endpunkten x, y; danach kann der Satz § 10, VII auch so gefaßt werden, daß sich in einem zusammenhängenden Gebiet je zwei Punkte x, y durch einen Weg $(x, \ldots, y)$ verbinden lassen.

Zwei Wege $(a_0, a_1, \ldots, a_m)$ und $(a_m, a_{m+1}, \ldots, a_{n-1}, a_0)$, die nur die Endpunkte a_0, a_m gemein haben, geben als Summe ein (einfaches, geschlossenes) Polygon

$$P = (a_0, a_1, \ldots, a_m, a_{m+1}, \ldots, a_{n-1}, a_0);$$

definiert man für alle ganzen Zahlen $a_i = a_k$, falls $i \equiv k \pmod{n}$, so ist auch

$$P = (a_i, a_{i+1}, \ldots, a_{i+n}).$$

Wir entnehmen der Elementargeometrie die Tatsache, daß jedes Polygon die Ebene in zwei zusammenhängende Gebiete J und A teilt, d. h. daß

$$E - P = J + A \quad \text{oder} \quad E = J + P + A;$$

das eine dieser Gebiete (J) ist beschränkt und heißt das Innere[1] von P, das andere (A) erstreckt sich ins Unendliche und heißt das Äußere von P. Überdies ist jeder Polygonpunkt Häufungspunkt sowohl von J als von A, woraus hervorgeht, daß das Polygon die gemeinsame Grenze beider Gebiete ist:

$$P = J_g = A_g.$$

Wir schreiben

$$E = M + N + P,$$

wenn wir uns die Entscheidung, ob $M = J$, $N = A$ oder umgekehrt $M = A$, $N = J$ sein soll, noch vorbehalten wollen.

---

[1] Abweichend von unserer sonstigen Bezeichnung, wonach P_i das Innere von P genannt wurde; hier ist $P_i = 0$.

Sind P, P' zwei Polygone,

$$E = J + P + A = J' + P' + A',$$

und sind P und P' fremd, so muß P als zusammenhängende Menge ganz in J' oder A' verlaufen und P' ganz in J oder A. Um alle vier zunächst denkbaren Fälle zugleich zu behandeln, schreiben wir

$$P \leqq M', \quad P' \leqq M.$$

Dann ist

$$N = \mathfrak{D}(N, M') + \mathfrak{D}(N, N'),$$
$$N' = \mathfrak{D}(M, N') + \mathfrak{D}(N, N').$$

Diese Zerlegungen müssen, weil N und N' zusammenhängend sind, einen verschwindenden Summanden haben, und zwar muß

$$\mathfrak{D}(N, N') = 0$$

sein; denn andernfalls wäre $N = N' = \mathfrak{D}(N, N')$ und $N_g = N_g'$, also $P = P'$, während die beiden Polygone doch verschieden sind. Also ist

$$N = \mathfrak{D}(N, M'), \quad N' = \mathfrak{D}(M, N')$$

oder

$$P + N \leqq M', \quad P' + N' \leqq M$$

und

$$E = P + P' + N + N' + \mathfrak{D}(M, M').$$

Dabei scheidet der Fall $N = A$, $N' = A'$ aus, weil $\mathfrak{D}(A, A')$ alle hinlänglich entfernten Punkte der Ebene enthält und demnach nicht verschwinden kann, und es bleiben die Fälle:

(1) $\quad P + J \leqq A', \; P' + J' \leqq A, \quad E = P + P' + J + J' + \mathfrak{D}(A, A')$

(2) $\quad P + J \leqq J', \; P' + A' \leqq A, \quad E = P + P' + J + A' + \mathfrak{D}(A, J')$

(3) $\quad P + A \leqq A', \; P' + J' \leqq J, \quad E = P + P' + A + J' + \mathfrak{D}(J, A')$,

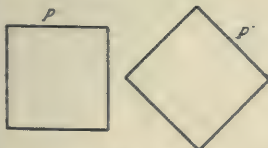


Fig. 13.



Fig. 14.

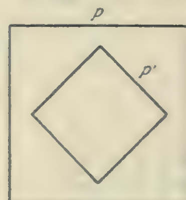


Fig. 15.

wo P und P' außerhalb von einander liegen, resp. P innerhalb P', resp. P' innerhalb P liegt.

Wir wollen ferner als evident annehmen, daß man ein das Polygon P umschließendes Polygon P' (Fall (2)) so zeichnen kann, daß die Entfernung $\overline{PP'}$ beider Polygone (S. 293) eine vorgegebene positive Zahl ϱ nicht überschreitet (Fig. 16). Gehört P einem

Gebiet G an und wählt man ϱ kleiner als den positiven Minimalabstand[1] der abgeschlossenen Mengen P und $E-G$, so verläuft auch P' noch in G, also $P' \leqq \mathfrak{D}(G, A)$. Ebenso kann man ein von P umschlossenes und zu P beliebig benachbartes Polygon angeben.

I. Jedes zusammenhängende Gebiet G wird durch ein darin verlaufendes Polygon P in zwei zusammenhängende Gebiete zerlegt.

Wir behaupten also, daß in der Formel

$$G = \mathfrak{D}(G, J) + \mathfrak{D}(G, P) + \mathfrak{D}(G, A) = J_0 + P + A_0$$

die beiden Gebiete J_0 und A_0 zusammenhängend sind. Zunächst sind beide von 0 verschieden, da jede Umgebung eines Punktes von P Punkte sowohl von J als von A enthält. Sind ferner (Fig. 17) a, b zwei Punkte von A_0, so verbinden wir sie in G durch einen Weg $W = (a, \ldots, b)$; sollte dieser P nicht treffen, also ganz in A oder A_0 verlaufen, so wäre unser Ziel schon erreicht. Andernfalls ziehen

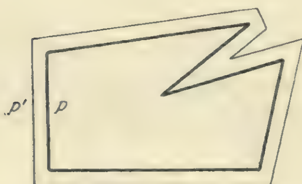


Fig. 16.

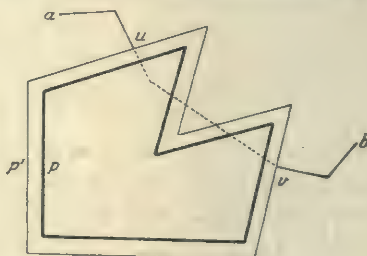


Fig. 17.

wir ein P umschließendes Polygon $P' \leqq A_0$ so nahe an P, daß a, b noch außerhalb P' (in A') liegen. W trifft dann $P \leqq J'$, also auch P', und die abgeschlossene Menge $\mathfrak{D}(W, P')$ hat auf dem von a nach b orientierten[2] Wege W genau wie auf der geradlinigen Strecke (S. 258) einen ersten Punkt u und einen letzten v $(a < u < v < b)$. Auf dem Wegstück $(a, \ldots, u)$ liegt kein Punkt von P' außer u, also kein Punkt von J' und keiner von $P + J \leqq J'$; es gehört ganz zu A und folglich zu $\mathfrak{D}(A, G) = A_0$. Das Gleiche gilt von $(v, \ldots, b)$, und wenn wir von u nach v, statt auf dem Wege W, auf der einen Hälfte des Polygons P' wandern, so haben wir a, b durch einen ganz zu A_0 gehörigen Weg verbunden. Damit ist der Zusammenhang von A_0 erwiesen, und dasselbe gilt von J_0.

[1] Ein solcher existiert zwischen zwei abgeschlossenen beschränkten Mengen, offenbar aber auch, wenn nur die eine dieser Mengen beschränkt ist.

[2] Man ordne jedem Punkt x von W in ersichtlicher Weise die Länge l_x des Wegstückes $(a, \ldots, x)$ zu und definiere $x \leqq y$ für $l_x \leqq l_y$.

Wir können dem Satz I noch folgende Form geben:

II. Ist G ein zusammenhängendes Gebiet, dessen Grenze H das Polygon P nicht trifft, so ist jedes der Gebiete $\mathfrak{D}(G, J)$ und $\mathfrak{D}(G, A)$, falls von Null verschieden, zusammenhängend.

Denn P gehört als zusammenhängende Menge entweder ganz zu G, und dann tritt I in Kraft; oder es gehört ganz zu $E-G$, und dann ist $G = \mathfrak{D}(G, J) + \mathfrak{D}(G, A)$, also einer der Summanden Null und der andere $= G$.

Der Satz II wieder zieht unmittelbar den folgenden nach sich:

III. Ist F eine abgeschlossene Menge, die das Polygon P nicht trifft, und werden zwei Punkte x, y weder durch F noch durch P getrennt, so werden sie auch durch $F+P$ nicht getrennt.

Denn x, y gehören einer Komponente G von $E-F$ an, und die Grenze H dieser Komponente liegt in F (S. 332), trifft also P nicht. Ebenso gehören x, y derselben Komponente M (J oder A) von $E-P$ an, gehören also dem zusammenhängenden Gebiet $\mathfrak{D}(G, M) \leqq E-(F+P)$ an und werden durch $F+P$ nicht getrennt.

Durch wiederholte Anwendung folgt, daß zwei Punkte auch durch $F+P_1+P_2 \ldots + P_n$ nicht getrennt werden, falls sie weder durch die abgeschlossene Menge F noch durch eins der Polygone $P_1, P_2, \ldots, P_n$ getrennt werden.

Um weiter zu gelangen, müssen wir das Polygon P des Satzes III durch eine zweite abgeschlossene Menge zu ersetzen suchen. Dazu dient die Approximation einer beschränkten Menge durch eine von Polygonen begrenzte Fläche.

Denken wir uns in der Ebene eine Schar von parallelen äquidistanten Geraden gezogen, und eine zweite dazu senkrechte Schar mit gleichem Abstand, wodurch ein quadratisches Gitter entsteht. Unter einem Quadrat Q_1 des Gitters verstehen wir die Quadratfläche (Inneres und Rand); es sei $Q = \mathfrak{S}(Q_1, Q_2, \ldots, Q_m)$ eine Summe von endlich vielen verschiedenen Quadraten. Dabei treffen wir folgende Bestimmung: wenn zwei nur mit einer Ecke zusammenstoßende Quadrate (wie 1, 1 in der Figur) zu Q gehören, so soll auch noch mindestens ein drittes Quadrat mit dieser Ecke (wie 2) zu Q gehören. Wenn Q diese Bedingung noch nicht erfüllt, so ist es möglich, eine endliche Anzahl weiterer Quadrate $Q_{m+1}, \ldots, Q_n$, von denen jedes mit einer Seite an eins der schon vorhandenen stößt, so hinzuzufügen, daß $Q^* = \mathfrak{S}(Q_1, Q_2, \ldots, Q_n)$ die Bedingung



Fig. 18.

erfüllt. Bezeichnen wir (Fig. 19) die Quadrate von Q mit 1 und ordnen sie in Zeilen von oben nach unten, indem wir nur diejenigen horizontalen Quadratreihen der Ebene betrachten, in denen wirklich Quadrate 1 stehen. Die oberste Zeile von Q lassen wir unverändert. Wenn in der zweiten Zeile ein Quadrat mit einem der ersten Zeile diagonal zusammenstößt, so fügen wir das beiden benachbarte Quadrat der zweiten Zeile hinzu, falls es nicht schon selber ein Quadrat 1 und falls nicht schon in der ersten Zeile ein zu jenen beiden benachbartes Quadrat vorhanden ist, und bezeichnen diese hinzugefügten Quadrate mit 2. (Eine solche Hinzufügung unterbleibt also sicher, wenn die zweite Zeile nicht direkt unter der ersten steht, sondern dazwischen eine horizontale Quadratreihe freibleibt.) Die beiden ersten Zeilen erfüllen dann die gestellte Bedingung; denn wenn ein Quadrat 2 mit einem der Quadrate 1 diagonal zusammenstößt, so ist ein zu beiden benachbartes Quadrat, nämlich das über 2 stehende Quadrat 1, immer schon vorhanden; außerdem hat

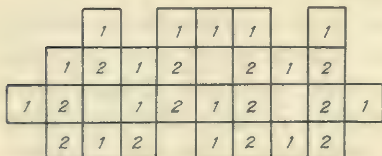


Fig. 19.

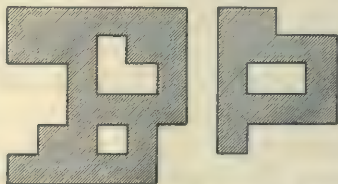


Fig. 20.

jedes Quadrat 2 als linken oder rechten Nachbarn ein Quadrat 1. Nunmehr wird die dritte Zeile mit Rücksicht auf die schon vervollständigte zweite Zeile ebenso behandelt, dann die vierte Zeile mit Rücksicht auf die vervollständigte dritte Zeile usw. Die Figur zeigt den Übergang von Q zu Q^* im Falle von vier Zeilen; die neu hinzugekommenen Quadrate 2 stoßen immer mit einer Seite an ein linkes oder rechtes Nachbarquadrat 1.

Sei nun $Q = \mathfrak{S}(Q_1, Q_2, \ldots, Q_m)$ eine Quadratsumme (Fig. 20) von der geforderten Eigenschaft; wir behaupten dann, daß der Rand (oder die Grenze) von Q aus einer endlichen Anzahl paarweise fremder Polygone besteht. Dies wird durch den Schluß von m auf $m+1$ bewiesen. Wir nehmen (Fig. 21—27) ein weiteres Quadrat Q_0 hinzu und bilden $Q' = \mathfrak{S}(Q, Q_0)$; die Ränder dieser Figuren seien R', R und $R_0 = (a, b, c, d, a)$. Die Menge $\mathfrak{D}(R, R_0)$, in der Q_0 an die schon vorhandene Menge Q stößt, ist entweder Null und dann ist $R' = R + R_0$; oder sie ist nicht Null und kann dann aus 1, 2, 3, 4 Seiten von R_0 bestehen (aber z. B. nicht aus einer Ecke allein). Stößt Q_0 in einer Seite (a, b) oder zwei anliegenden Seiten (a, b, c) oder drei Seiten

(a, b, c, d) an Q, so werden diese Teile von R beim Übergang zu R' durch (a, d, c, b) resp. (a, d, c) resp. (a, d) ersetzt, also eins der vorhandenen Polygone in ein anderes verwandelt.



Fig. 21.

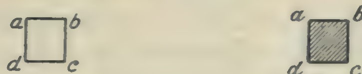


Fig. 22.

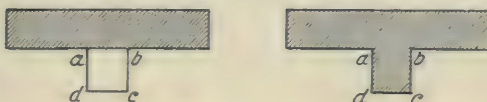


Fig. 23.

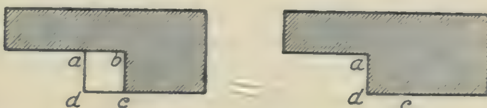


Fig. 24.



Fig. 25.



Fig. 26.



Fig. 27.



Stößt Q_0 mit zwei gegenüberliegenden Seiten (a, b), (c, d) an Q, so fallen in der neuen Figur Q' diese Randstrecken fort und dafür treten (a, d) und (b, c) auf, wodurch entweder zwei bestehende Poly-

gone in eins verwandelt oder umgekehrt aus einem zwei gemacht werden. Stößt endlich Q_0 mit allen vier Seiten an Q, so ist $R' = R - R_0$; das Quadrat R_0 kommt in Wegfall.

Sei nun A irgend eine beschränkte Menge: es gibt dann eine Quadratsumme Q der eben betrachteten Art, die die Menge A in ihrem Innern einschließt ($A \leqq Q_i$) und von ihr eine beliebig kleine Entfernung $\overline{AQ}$ hat (§ 6). Wir überziehen nämlich die Ebene mit einem quadratischen Gitter von der Seitenlänge σ, und es seien $Q_1, Q_2, \ldots, Q_l$ diejenigen Gitterquadrate, die einen Punkt von A enthalten (wie verabredet, im Innern oder auf dem Rande); die Summe dieser Quadrate enthält jeden Punkt a und zwar bereits als inneren Punkt, denn wenn a auf einer Seite oder in einer Ecke eines Quadrats liegen sollte, so gehört er zwei oder vier Quadraten an und ist innerer Punkt des betreffenden Rechtecks oder Doppelquadrats. Durch eventuelle Hinzunahme weiterer Quadrate, wie oben gezeigt, erhalten wir dann eine Quadratsumme $Q = \mathfrak{S}(Q_1, Q_2, \ldots, Q_m)$, die der Bedingung für diagonal zusammenstoßende Quadrate genügt und deren Rand also eine Summe paarweise fremder Polygone ist: $Q_r = P_1 + P_2 + \cdots + P_n$. Dabei enthält jedes der Quadrate von Q entweder selbst einen Punkt von A oder stößt mit einer Seite an ein solches an, es gibt also zu jedem Punkt q einen Punkt a mit

$$\overline{aq} \leqq \sqrt{5}\,\sigma < 3\sigma.$$

Also ist (S. 293) $\overrightarrow{AQ} < 3\sigma$ und $\overrightarrow{QA} = 0$ (wegen $A \leqq Q$), also $\overline{AQ} < 3\sigma$.

Ist B eine weitere Menge mit positiver unterer Entfernung $\delta(A, B)$ und wählt man σ hinreichend klein, etwa $3\sigma < \delta(A, B)$, so haben Q und B keinen Punkt gemein. Zu zwei Mengen A, B mit positiver unterer Entfernung, von denen A beschränkt ist, kann man also eine Quadratsumme Q angeben, die A einschließt und B ausschließt, und deren Rand eine Summe paarweise fremder Polygone ist:

$$A \leqq Q_i, \quad B \leqq E - Q, \quad Q_r = P_1 + P_2 + \cdots + P_n.$$

Nunmehr können wir den Satz beweisen:

IV. Sind F_1, F_2 zwei fremde abgeschlossene beschränkte[1] Mengen, und werden zwei Punkte weder durch F_1 noch durch F_2 getrennt, so werden sie auch durch $F_1 + F_2$ nicht getrennt.

Da x, y durch F_1 nicht getrennt werden, so gibt es (Fig. 28) einen

[1] Es genügt, daß die eine Menge (beim Beweise F_1) beschränkt sei; auch einige der folgenden Sätze lassen sich in diesem Sinne etwas verallgemeinern. Weiter kann man nicht gehen, ohne statt der Wege und Polygone kompliziertere Figuren (solche mit unendlich vielen Strecken) zu benutzen.

sie verbindenden Weg W_2, der F_1 nicht trifft. Wir bestimmen eine Quadratsumme Q, die F_1 einschließt und $\mathfrak{S}(F_2, W_2)$ ausschließt:

$$F_1 \leqq Q_i, \quad \mathfrak{S}(F_2, W_2) \leqq E - Q.$$

x, y selbst gehören zu $E - Q$. Da W_2 die Menge Q nicht trifft, so werden x, y durch Q oder durch $Q_r = P_1 + P_2 + \ldots + P_n$ nicht getrennt, also nach III auch nicht durch $F_2 + Q_r$. Demnach gibt es einen x, y verbindenden Weg W, der F_2 und Q_r nicht trifft. Dieser kann aber auch F_1 nicht treffen, da er sonst einen Punkt x von $E - Q$ mit einem Punkte von $F_1 \leqq Q$ verbinden würde, also auch die Grenze $Q_g = Q_r$ treffen müßte. Da W also $F_1 + F_2$ nicht trifft, so werden die Punkte durch diese Menge nicht getrennt.

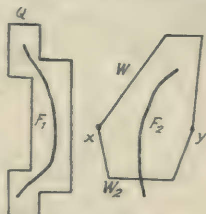


Fig. 28.

Von diesem Satz IV, der eine Verallgemeinerung von III ist, kann man zu einer entsprechenden Verallgemeinerung von II zurückgelangen:

V. Zwei zusammenhängende Gebiete, deren beschränkte Grenzen einander nicht treffen, haben einen zusammenhängenden oder verschwindenden Durchschnitt.

Die Gebiete seien G_1, G_2 mit den Grenzen H_1, H_2. Wenn $G = \mathfrak{D}(G_1, G_2)$ nicht Null ist, so werden zwei seiner Punkte als Punkte von G_1 durch H_1 nicht getrennt, als Punkte von G_2 nicht durch H_2, also auch nicht durch $H_1 + H_2$; sie lassen sich also durch einen Weg W verbinden, der H_1 und H_2 nicht trifft. Dieser Weg muß dann ganz in G_1 verlaufen, ebenso ganz in G_2, also ganz in G; G ist zusammenhängend.

Umgekehrt folgt aus V ebenso unmittelbar wieder IV: wenn zwei Punkte durch F_1 und F_2 nicht getrennt werden, so gehören sie einer Komponente G_1 von $E - F_1$ mit der Grenze $H_1 \leqq F_1$ und einer Komponente G_2 von $E - F_2$ mit der Grenze $H_2 \leqq F_2$, also nach V dem zusammenhängenden Gebiet $G = \mathfrak{D}(G_1, G_2) \leqq E - (F_1 + F_2)$ an und werden durch $F_1 + F_2$ nicht getrennt.

Beide Sätze übertragen sich unmittelbar von zwei auf endlich viele abgeschlossene Mengen resp. Gebiete. In Verbindung mit § 10, VIII gestatten sie auch eine Ausdehnung auf das Abzählbare, der wir folgende Form geben:

VI. Wenn zwei Punkte durch die beschränkte abgeschlossene Menge F getrennt werden, so werden sie auch durch eine Komponente von F getrennt.

Denn wenn $F = F_0$ nicht zusammenhängend, also in zwei abgeschlossene Summanden zerlegbar ist, so muß einer davon, etwa

$F_1 < F_0$, die gegebenen Punkte trennen; ist F_1 noch nicht zusammenhängend, so werden die Punkte wieder durch eine abgeschlossene Menge $F_2 < F_1$ getrennt usw. Gelangt man noch nicht für endliches ν zu einer zusammenhängenden Menge F_ν, so trennt auch $F_\omega = \mathfrak{D}(F_0, F_1, F_2, \ldots)$ die Punkte noch, und man kann das Verfahren fortsetzen. Auf Grund von § 4 gelangt man nach einer höchstens abzählbaren Menge von Schritten zu einer letzten Menge F_η, die noch die Punkte trennt, zu der es aber keine kleinere Menge $F_{\eta+1} < F_\eta$ mit der gleichen Eigenschaft gibt; dann ist also F_η zusammenhängend, und die Punkte werden durch F_η oder durch die Komponente $\geqq F_\eta$ getrennt. Wenn man die Zerlegung in ϱ-Komponenten benutzt (§ 6), kann man übrigens erreichen, daß das Verfahren spätestens bei F_ω abbricht.

Beispiel: das Komplement einer beschränkten punkthaften (S. 322) Menge C ist zusammenhängend. Denn wäre $E - C = A + B$ in zwei relativ abgeschlossene Mengen zerlegbar, so gäbe es nach S. 335 eine abgeschlossene Menge $F \leqq C$, die zwei Punkte von A und B trennt; es müßte dann eine Komponente von F, also eine Menge aus einem einzelnen Punkt c, die beiden Punkte trennen, was wegen des Zusammenhanges von $E - \{c\}$ aber nicht der Fall ist.

In den weiteren Anwendungen spielt folgende Spezialisierung der bisherigen Sätze die Hauptrolle.

VII. Zwei Punkte, die verschiedenen Komponenten der beschränkten abgeschlossenen Menge H angehören, lassen sich durch ein Polygon P trennen, das H nicht trifft. Ist H insbesondere Grenze des zusammenhängenden Gebietes G, so liegt das Polygon P im Gebiet G.

Die beiden Punkte seien h_1, h_2; da bei einer beschränkten abgeschlossenen Menge Komponenten und Quasikomponenten identisch sind (S. 303), so gibt es eine Zerlegung $H = H_1 + H_2$ in abgeschlossene Mengen, bei der h_1 zu H_1 und h_2 zu H_2 gehört. Bestimmen wir eine Quadratsumme Q mit

$$H_1 \leqq Q_i, \quad H_2 \leqq E - Q,$$

so werden h_1, h_2 durch $Q_r = P_1 + P_2 + \ldots + P_n$, also nach III durch eins dieser Polygone getrennt (das natürlich nicht für alle diese Punktpaare dasselbe zu sein braucht). — Um den zweiten Teil des Satzes zu beweisen, bemerken wir, daß sowohl innerhalb wie außerhalb des Polygons, nämlich in hinreichend kleinen Umgebungen von h_1 und h_2, Punkte von G liegen; da das Gebiet G zusammenhängend ist, hat es auch mit P Punkte gemein. Das Polygon P trifft also G; wenn es auch $E - G$ träfe, müßte es als zusammenhängende Menge die Grenze H treffen; da dies nicht der Fall ist, liegt P ganz in G.

VIII. Wenn die beschränkte Grenze H eines zusammenhängenden Gebietes G einer zusammenhängenden Teilmenge von $F = E - G$ angehört, so ist sie selbst zusammenhängend.[1]

Es sei $H \leqq C \leqq F$, C zusammenhängend. Wäre H nicht zusammenhängend, so gäbe es nach VII ein in G verlaufendes Polygon P, das innerhalb wie außerhalb Punkte von H, also auch von C enthielte, und C müßte das Polygon treffen, während doch C und G keinen Punkt gemein haben.

Spezialfälle dieses Satzes sind:

IX. Ist F abgeschlossen und zusammenhängend, G_1 eine Komponente von $G = E - F$ mit beschränkter Grenze H_1, so ist H_1 zusammenhängend.

Denn es ist $H_1 \leqq H \leqq F \leqq F_1$ und H_1 als Teilmenge von F zusammenhängend.[2] Vgl. Fig. 29.

X. Ist G ein zusammenhängendes Gebiet, F_1 eine Komponente von $F = E - G$ mit beschränktem Rand H_1, so ist H_1 zusammenhängend.

Sei nämlich $F = F_1 + F_2 + F_3 + \ldots$ die Zerlegung von F in seine (endlich, abzählbar oder unabzählbar vielen) Komponenten. Da eine zusammenhängende Menge nach Hinzufügung beliebig vieler Komponenten ihres Komplements zusammenhängend bleibt (S. 331), so ist $G_1 = G + F_2 + F_3 + \ldots$ zusammenhängend, nach VIII ist also H_1 als Teilmenge der zusammenhängenden Menge F_1 selbst zusammenhängend.

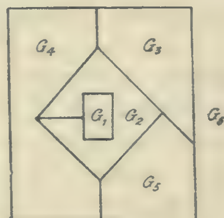


Fig. 29.

Wenn alle H_n beschränkt und daher zusammenhängend sind, so sind sie die Komponenten von H, da eine zusammenhängende Teilmenge von H auch eine von F ist, also nicht mit zwei verschiedenen F_n oder H_n Punkte gemein haben kann. Das gibt insbesondere den Satz:

XI. Ist G ein zusammenhängendes Gebiet mit beschränkter Grenze H, so hat diese ebenso viele Komponenten wie $F = E - G$.

Wir schließen diese Reihe von Überlegungen mit dem Satz:

---

[1] Wir erinnern beiläufig daran, daß eine Gebietsgrenze Randmenge und nirgendsdicht, eine zusammenhängende Gebietsgrenze perfekt ist (falls sie nicht aus einem Punkte besteht).

[2] Wenn nichts anderes ersichtlich ist, bedeute immer F das Komplement, H die Grenze des Gebietes G; ebenso mit Indices oder Akzenten.

XII. Sind G_1, G_2 zwei fremde zusammenhängende Gebiete mit den Grenzen H_1, H_2, ferner H_1 beschränkt und in H_2 enthalten, so ist H_1 zusammenhängend.

Denn $G_2 + H_1$ ist als Menge zwischen G_2 und $G_2 + H_2$ zusammenhängend, $H_1 \leqq G_2 + H_1 \leqq E - G_1$, also H_1 zusammenhängend. Insbesondere ist, für $H_1 = H_2$, die gemeinsame Grenze zweier fremder zusammenhängender Gebiete stets zusammenhängend.

Eine besondere Rolle spielen diejenigen beschränkten abgeschlossenen Mengen H, die die gemeinsame Grenze aller Komponenten von $E - H$ sind, also

$$E - H = G_1 + G_2 + \dots, \quad H = H_1 = H_2 = \dots;$$

eine solche Menge H ist nach XII zusammenhängend, vorausgesetzt, daß $E - H$ mindestens zwei Komponenten hat.[1] Falls genau zwei Komponenten auftreten, $E - H$ also in zwei zusammenhängende Gebiete mit der gemeinsamen Grenze H zerfällt, wird H (nach A. Schoenflies) eine geschlossene Kurve genannt; eine geschlossene Kurve ist jedenfalls eine beschränkte, perfekte, zusammenhängende Menge ohne innere Punkte. Die einfachsten geschlossenen Kurven sind die Polygone ($E - P = J + A$, $P = J_g = A_g$) und, nach dem unten zu beweisenden Jordanschen Satz, deren umkehrbar eindeutige, stetige Bilder.

Es ist leicht einzusehen, daß sich bezüglich der Zerlegung der Ebene in Gebiete die geschlossenen Kurven genau wie Polygone verhalten. Ist P eine geschlossene Kurve, so ist das Äußere eines Kreises, in dessen Innerem P liegt, zusammenhängend und gehört einer der Komponenten von $E - P$ an; wir können also wieder $E - P = A + J$ setzen, wo A alle hinlänglich entfernten Punkte der Ebene enthält und J beschränkt ist (A das Äußere, J das Innere von P). Über die gegenseitige Lage zweier fremder geschlossener Kurven gelten dann wörtlich dieselben Betrachtungen (S. 337) wie über Polygone; der Satz I gilt auf Grund von V auch, wenn P durch eine geschlossene Kurve ersetzt wird. Eine Menge paarweise fremder geschlossener Kurven $P, P', \dots$, von denen keine die übrigen trennt, ist höchstens abzählbar: denn liegen alle Kurven außer P in M ($M = J$ oder A), alle außer P' in M' usw., und setzt man wieder $E = M + P + N = M' + P' + N' = \dots$, so ist $\mathfrak{D}(N, N') = 0$ usw., die Gebiete $N, N', \dots$ sind also paarweise fremd und ihre Menge höchstens abzählbar.

Ist H eine beschränkte abgeschlossene Randmenge, deren

[1] Daß unter den angegebenen Voraussetzungen auch mehr als zwei Komponenten vorkommen können, ist der gewöhnlichen Anschauung schwer vorstellbar; L. E. J. Brouwer hat ein Beispiel dafür gegeben.

Komplement $E - H = G_1 + G_2$ in zwei Komponenten zerfällt, so braucht $H = (G_1 + G_2)_g = \mathfrak{S}(H_1, H_2)$ keine geschlossene Kurve zu sein; wohl aber ist $\mathfrak{H} = \mathfrak{D}(H_1, H_2)$ eine solche. Denn setzt man

$$\mathfrak{G}_1 = G_1 + (H_1 - \mathfrak{H}), \quad \mathfrak{G}_2 = G_2 + (H_2 - \mathfrak{H}),$$
$$H = (H_1 - \mathfrak{H}) + (H_2 - \mathfrak{H}) + \mathfrak{H},$$

so ist

$$E = G_1 + G_2 + H = \mathfrak{G}_1 + \mathfrak{G}_2 + \mathfrak{H}$$
$$= \mathfrak{G}_1 + G_2 + H_2 = \mathfrak{G}_2 + G_1 + H_1.$$

Da $G_1 + H_1 = G_{1\alpha}$ abgeschlossen ist, so ist $\mathfrak{G}_2$ ein Gebiet; ebenso $\mathfrak{G}_1$; beide sind fremd. Ferner ist $G_1 \leqq \mathfrak{G}_1 \leqq G_{1\alpha}$, also $\mathfrak{G}_1$ (und $\mathfrak{G}_2$) zusammenhängend, und $\mathfrak{G}_{1\alpha} = G_{1\alpha} = G_1 + H_1 = \mathfrak{G}_1 + \mathfrak{H}$, also $\mathfrak{H}$ die Grenze von $\mathfrak{G}_1$ (und $\mathfrak{G}_2$). Danach zerfällt $E - \mathfrak{H} = \mathfrak{G}_1 + \mathfrak{G}_2$ in zwei Komponenten mit der gemeinsamen Grenze $\mathfrak{H}$, $\mathfrak{H}$ ist eine geschlossene Kurve.

Beispiel: H sei die Streckensumme der nebenstehenden Figur; $\mathfrak{H}$ ist das in H enthaltene Polygon, $\mathfrak{G}_1$ und $\mathfrak{G}_2$ dessen Äußeres und Inneres; die Strecken, die vorher zur Grenze nur eines Gebietes gehörten, haben sich mit diesem vereinigt.



Fig. 30.

Ein Punkt h der Grenze H des zusammenhängenden Gebietes G heiße von G aus **geradlinig erreichbar**, wenn man ihn mit einem Punkte g von G durch einen Weg $(g, \ldots, h)$ verbinden kann, der bis auf den Punkt h in G liegt. Offenbar kann man ihn dann mit jedem Punkt g' von G so verbinden, indem man g' mit g durch einen Weg verbindet und den Streckenzug $(g', \ldots, g, \ldots, h)$ zu einem Wege reduziert (S. 336). Z. B. sind bei einem Polygon oder Kreise alle Punkte sowohl von J wie von A aus geradlinig erreichbar; dagegen begrenzt eine Gerade und ein sie berührender Kreis ein Gebiet, von dem aus der Berührungspunkt nicht geradlinig erreichbar ist.

Die Menge der geradlinig erreichbaren Punkte ist in H dicht. Denn ist h ein beliebiger Punkt der Grenze, so gibt es für ein beliebig kleines positives ϱ einen Punkt g des Gebietes mit $\overline{hg} < \varrho$; ist dann h' der dem Punkte g nächste Punkt der (abgeschlossenen) Menge H, so ist $\overline{gh'} \leqq \overline{gh} < \varrho$, also $\overline{hh'} \leqq \overline{hg} + \overline{gh'} < 2\varrho$. Da nun die Strecke (g, h') keinen Punkt von H außer h' enthält und sonst in G liegt, so ist h' geradlinig erreichbar, und es befindet sich in jeder Umgebung eines Punktes h ein geradlinig erreichbarer Punkt h'.

Man nennt ferner h von G aus **schlechthin erreichbar**, wenn man ihn mit einem Punkte g von G durch eine **abgeschlossene zusammenhängende** Menge verbinden kann, die bis auf h in G

liegt. Zu den erreichbaren Punkten gehören natürlich die geradlinig erreichbaren, und die Menge der erreichbaren Punkte ist um so mehr in H dicht. Das einfachste Beispiel einer Gebietsgrenze H mit nicht sämtlich erreichbaren Punkten liefert die Abszissenachse samt den in den Punkten $x = 0, \pm 1, \pm \frac{1}{2}, \pm \frac{1}{4}, \pm \frac{1}{8}, \ldots$ nach oben errichteten Loten von der Höhe 1. Die Punkte h des zu $x = 0$ gehörigen Lotes (h_0, h_1), mit Ausnahme des oberen Endpunktes h_1, sind von dem durch H begrenzten Gebiet G (in der oberen Halbebene) aus nicht erreichbar. Denn denkt man sich die gerade Linie $y = 1$ gezogen, so entstehen unendlich viele Rechtecke, und die zusammenhängende Menge C, die h zum Häufungspunkt hat und bis auf h zu G gehört, muß innere Punkte von unendlich vielen dieser Rechtecke enthalten, muß also, um von einem ins andere zu gelangen, die Gerade $y = 1$ in unendlich vielen Punkten treffen, die den Punkt h_1 zum Häufungspunkt haben. Da C außerdem als abgeschlossen vorausgesetzt war, so müßte sie auch den genannten Punkt h_1 (ebenso jeden Punkt der Vertikalstrecke zwischen h und h_1) enthalten, was gegen die Forderung verstößt.

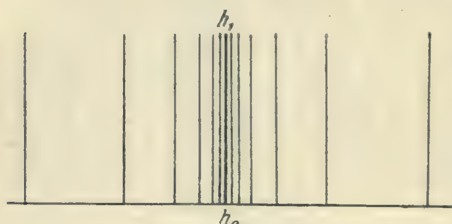


Fig. 31.

Unter einem Querschnitt des zusammenhängenden Gebietes G versteht man einen Weg $W = (a, \ldots, b)$, der bis auf die Endpunkte in G liegt, während diese zu $F = E - G$, also offenbar zur Gebietsgrenze H gehören. Den Querschnitt ohne die Endpunkte bezeichnen wir mit

$$V = \mathfrak{D}(W, G) = W - \{a, b\}$$

und die nach Tilgung von V verbleibende Menge mit

$$G' = G - V = \mathfrak{D}(G, E - W);$$

sie ist wieder ein Gebiet.

Offenbar ist G' in G dicht, also

$$G'_\alpha = G_\alpha = G + H = G' + H + V,$$

demnach die Grenze des neuen Gebiets $H' = H + V$; sie besteht aus der alten Grenze und dem Querschnitt.

Die Hauptfrage ist nun, ob G durch den Querschnitt zerfällt oder nicht, d. h. ob G' unzusammenhängend oder zusammenhängend ist. Wir zeigen zunächst, daß G' höchstens in zwei Komponenten zerfallen kann. Wir können jeden Punkt x von G' mit irgend einem Punkt von V durch einen in G liegenden Weg W_x verbinden; orientieren wir W_x von x

aus, so hat auf W_x die abgeschlossene Menge $\mathfrak{D}(W_x, W) = \mathfrak{D}(W_x, V)$ einen ersten Punkt v, und je nachdem der Weg $(x, \ldots, v)$ sich dem (ebenfalls, etwa von a nach b orientierten) Wege W von dessen linker oder rechter Seite nähert, bezeichnen wir x als einen Punkt x_λ oder x_ϱ, wobei aber x je nach dem gewählten Wege W_x vielleicht zu beiden Klassen gehören kann. Zwei Punkte x_λ, y_λ können dann, indem man ein Stück des Weges W durch ein benachbartes ersetzt (vgl. die analoge Bemerkung über benachbarte Polygone S. 337), auch durch einen Weg in G verbunden werden, der W nicht trifft, also in G' verläuft; und umgekehrt ist ein Punkt y, der mit x_λ in G' verbunden werden kann, ein Punkt y_λ. Die Menge G_λ der Punkte x_λ ist also eine Komponente von G', ebenso die Menge G_ϱ der Punkte x_ϱ, und es ist $G' = \mathfrak{S}(G_\lambda, G_\varrho)$; folglich ist entweder



Fig. 32.

$$G' = G_\lambda = G_\varrho$$

(wenn es Punkte x gibt, die gleichzeitig mit der linken und rechten Seite von W verbunden werden können), oder

$$G' = G_\lambda + G_\varrho.$$

Für die Grenzen H_λ, H_ϱ gelten im zweiten Fall (übrigens auch im ersten) die Formeln

$$\mathfrak{S}(H_\lambda, H_\varrho) = H' = H + V = \mathfrak{S}(H, W),$$

$$\mathfrak{D}(H_\lambda, H_\varrho) \geqq W,$$

deren zweite besagt, daß die Punkte des Querschnitts Häufungspunkte sowohl der Punkte x_λ als auch der Punkte x_ϱ sind.

Wir zeigen weiter, indem wir die Grenze H als beschränkt annehmen, daß der erste oder zweite Fall eintritt, je nachdem die Endpunkte a, b des Querschnitts zu verschiedenen oder zu derselben **Komponente** von H gehören.

Gehören a, b zu verschiedenen Komponenten, so lassen sie sich nach VII durch ein in G verlaufendes Polygon P trennen. Der Weg W, der einen Punkt innerhalb P mit einem Punkt außerhalb verbindet, muß dann P treffen, und eine einfache (den Elementen der „Charakteristikentheorie“ angehörige) Überlegung zeigt, daß P die linke mit der rechten Seite von W verbindet, so daß in diesem Fall $G' = G_\lambda = G_\varrho$ zusammenhängend ist. Man kann nämlich, indem man eventuell P durch ein benachbartes Polygon ersetzt, erreichen, daß W und P keine Strecke gemeinsam haben und sich an jedem ihrer Schnittpunkte kreuzen, d. h. W von der einen auf die andere

Seite von P und P von der einen auf die andere Seite von W tritt. Die Anzahl n der Schnittpunkte ist dann ungerade; auf dem mit einer Umlaufrichtung versehenen Polygon geordnet, seien $p_1, p_2, \ldots, p_n$ diese Schnittpunkte. Von den Wegen $P_{12} = (p_1, \ldots, p_2)$, $P_{23}, \ldots, P_{n1}$, die hierdurch auf dem Polygon bestimmt werden, muß mindestens einer von der linken auf die rechte Seite von W führen oder umgekehrt; denn andernfalls würde, wenn etwa P_{12} von λ nach λ führt, P_{23} von ϱ nach ϱ, P_{34} wieder von λ nach λ, schließlich P_{n1} von λ nach λ und dann wieder P_{12} von ϱ nach ϱ führen, was einen Widerspruch ergibt. Wenn also etwa P_{12} von λ nach ϱ führt, so ist P_{12}, ohne die Endpunkte, eine in G' verlaufende zusammenhängende Menge, die ein x_λ mit einem x_ϱ verbindet. Sollte nur ein Schnittpunkt p_1 vorhanden sein, so ist $P - \{p_1\}$ eine solche Menge.

Umgekehrt aber: ist G' zusammenhängend, so lassen sich zwei in hinlänglicher Nähe eines Punktes von V gewählte Punkte x_λ, x_ϱ durch einen Weg in G' und überdies direkt mit Überschreitung von V verbinden, wodurch nach eventueller Ausscheidung entbehrlicher Strecken ein Polygon P in G entsteht, das von W einmal gekreuzt wird, also die Punkte a, b trennt. Danach existiert eine Zerlegung $H = \mathfrak{D}(H, J) + \mathfrak{D}(H, A) = H_a + H_b$ in zwei abgeschlossene Mengen mit Separation von a und b, diese Punkte gehören also verschiedenen Komponenten von H an.

Wir haben also den Satz bewiesen:

XIII. **Ein zusammenhängendes Gebiet mit beschränkter Grenze H wird durch einen Querschnitt $(a, \ldots, b)$ in zwei zusammenhängende Gebiete zerlegt, wenn die Endpunkte a, b derselben Komponente von H angehören; es bleibt zusammenhängend, wenn a, b verschiedenen Komponenten von H angehören.**

Betrachten wir den zweiten Fall noch etwas genauer, wobei wir zunächst konstatieren, daß es bei einem Gebiet mit unzusammenhängender Grenze solche Querschnitte wirklich gibt: ist $H = H_1 + H_2$ eine Spaltung in zwei abgeschlossene Mengen, so muß die Menge K der geradlinig erreichbaren Punkte, weil sie in H dicht ist, Punkte sowohl mit H_1 als mit H_2 gemein haben (für $K \leqq H_1$ wäre $K_\alpha \leqq H_1$, während $K_\alpha = H$ ist); zwei solche gehören verschiedenen Komponenten an und lassen sich durch einen Querschnitt verbinden. Sei nun

$$H = A + B + C + D + \ldots$$

die Zerlegung von H in Komponenten, und der Querschnitt $W = (a, \ldots, b)$, dessen Endpunkte auf A und B liegen, verwandle G in das noch zusammenhängende Gebiet G' mit der Grenze

$$H' = H + V = (A + V + B) + C + D + \ldots$$

Es ist leicht einzusehen, daß die zusammenhängenden Mengen $A + V + B = \mathfrak{S}(A, W, B)$, $C, D, \ldots$ die Komponenten von H' sind.[1] Denn es gibt eine Zerlegung $H = H_a + H_b + H_c + H_d$ in abgeschlossene Mengen, die resp. A, B, C, D als Teilmengen enthalten, und dann zeigt die Zerlegung $H' = \mathfrak{S}(H_a + H_b, W) + H_c + H_d$, daß $A + V + B, C, D$ verschiedenen Komponenten von H' angehören, also selbst Komponenten sind. Durch den Querschnitt werden also zwei Komponenten von H zu einer von H' vereinigt, während die übrigen bleiben; die Komponentenzahl von H' ist um eine Einheit kleiner als die von H.

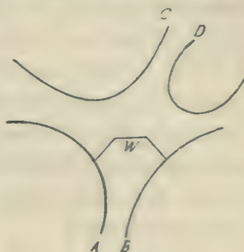


Fig. 33.

Wir wollen ein zusammenhängendes Gebiet mit beschränkter Grenze[2] *k-fach zusammenhängend* nennen, wenn seine Grenze (oder, nach XI, sein Komplement) k Komponenten hat; schließen wir die ganze Ebene aus, so ist nach S. 321 k entweder eine natürliche Zahl oder $\aleph_0$ oder $\aleph$, die Mächtigkeit des Kontinuums. Man kann also ein k-fach zusammenhängendes Gebiet G für $k > 1$ durch einen geeigneten (zwei verschiedene Komponenten der Grenze H verbindenden) Querschnitt W in ein $(k-1)$-fach zusammenhängendes Gebiet G' verwandeln, wobei, für $k = \aleph_0$ oder $k = \aleph$, natürlich $k - 1 = k$ zu setzen ist. Ist noch $k - 1 > 1$, so kann man G' wieder durch einen geeigneten Querschnitt W' (dessen Endpunkte verschiedenen Komponenten von H' angehören; der eine kann auch auf dem vorigen Querschnitt W liegen) in ein $(k-2)$-fach zusammenhängendes Gebiet G'', dieses eventuell wieder durch einen geeigneten Querschnitt W'' in ein $(k-3)$-fach zusammenhängendes Gebiet G''' verwandeln usw. Ist k eine natürliche Zahl, so lassen sich also sukzessiv $k-1$ Querschnitte ziehen, ohne, wie man sagt, das Gebiet zu zerstückeln, aber nicht mehr als $k-1$, denn das jetzt erhaltene Gebiet G^{k-1} ist einfach zusammenhängend und wird durch jeden weiteren Querschnitt unzusammenhängend. Ist k hingegen unendlich, so lassen sich beliebig viele sukzessive Querschnitte $W, W', W'', \ldots$ ziehen, ohne G zu zerstückeln; die verbleibenden Gebiete $G', G'', G''', \ldots$ sind immer noch k-fach zusammenhängend. Es wäre aber falsch, zu sagen, daß man unendlich viele Querschnitte ziehen kann,

---

[1] Der Beweis ist nur für unendliche Komponentenzahl erforderlich. Wir erinnern nochmals an die hier gültige Identität zwischen Komponenten und Quasikomponenten.

[2] Diese Definition ist hier bequemer; im Sinne der Invarianz bei eineindeutiger stetiger Abbildung (Kap. IX) ist bei unbeschränkten Gebieten eine andere Verabredung vorzuziehen.

ohne G zu zerstückeln, denn die schließlich übrigbleibende Menge $\mathfrak{D}(G, G', G'', \ldots)$ braucht weder ein Gebiet noch zusammenhängend zu sein.

Beispiel: Das nicht schraffierte Gebiet (Fig. 34) im Innern des großen Quadrats hat eine Grenze mit 4 Komponenten, ist 4-fach zusammenhängend; die punktierten Linien geben drei Querschnitte, die es in ein einfach zusammenhängendes verwandeln.

Ein einfach zusammenhängendes Gebiet G (also mit beschränkter zusammenhängender Grenze H) wird durch jeden Querschnitt zerstückelt[1]; dabei sind die entstehenden Gebiete G_λ, G_ϱ ebenfalls einfach zusammenhängend. In der Tat ist, nach VIII, H_λ als Teilmenge der zusammenhängenden Menge $H' = \mathfrak{S}(H, W) \leqq E - G_\lambda$ zusammenhängend.

Es seien A und B zwei fremde, beschränkte, abgeschlossene, zusammenhängende Mengen und

$$G = E - (A + B) = G_0 + G_1 + G_2 + \ldots$$

die Zerlegung des Komplementärgebiets in Komponenten. Für die

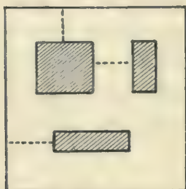


Fig. 34.

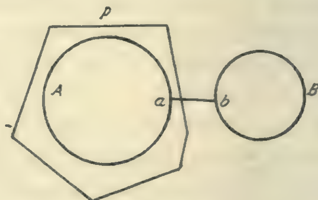


Fig. 35.

Grenzen gilt $H_n \leqq H = (A + B)_r$; H_n liegt in $A + B$. Zieht man nun (Fig. 35) die kürzeste Strecke (a, b) zwischen A und B, so liegt diese bis auf ihre Endpunkte in einer Komponente von G, etwa G_0, und die Grenze H_0 trifft also sowohl A als auch B. Wir behaupten, daß nur diese eine Komponente G_0 (die wir das Zwischengebiet von A und B nennen wollen) die genannte Eigenschaft hat, während die Grenzen H_n der übrigen G_n entweder ganz in A oder ganz in B liegen. Ziehen wir nämlich ein Polygon P, das, ohne $A + B$ zu treffen, zwei Punkte dieser Menge trennt, so liegt die eine der Mengen A, B ganz innerhalb, die andere ganz außerhalb P; P trifft also die Strecke (a, b), trifft mithin G_0 und liegt ganz in G_0. Jede andere Komponente G_n liegt dann ganz außerhalb oder ganz inner-

[1] Sollte H nur aus einem Punkte bestehen, G also die ganze Ebene mit Ausschluß eines Punktes sein, so existiert kein Querschnitt im bisherigen Sinne, d. h. mit zwei verschiedenen Endpunkten; diesen trivialen Fall schließen wir aus.

halb des Polygons und ihre Grenze H_n kann entweder nur mit A oder nur mit B Punkte gemein haben.

Das Zwischengebiet G_0 ist zweifach zusammenhängend. Denn sind $G_1, G_3, \ldots$ diejenigen G_n, deren Grenzen in A liegen, und $G_2, G_4, \ldots$ diejenigen, deren Grenzen in B liegen, so ist

$$A + G_1 = \mathfrak{S}(A, G_{1\alpha})$$

zusammenhängend, ebenso

$$A + G_1 + G_3 + \ldots = \mathfrak{S}(A + G_1, A + G_3, \ldots).$$

ebenso $B + G_2 + G_4 + \ldots$; die Summe der beiden letzten Mengen, $E - G_0$, hat also höchstens zwei Komponenten. Nach XI hat auch H_0 höchstens zwei Komponenten; da H_0 nicht zusammenhängend ist, genau zwei.[1]

Denken wir uns nun eine endliche Anzahl von Wegen $W_i = (a_i, \ldots, b_i) = \{a_i, b_i\} + V_i$, deren Endpunkte auf A, B liegen, während die V_i weder einander noch $A + B$ treffen ($i = 1, 2, \ldots$); die paarweise fremden V_i verlaufen also im Zwischengebiet von A und B, das wir jetzt G nennen wollen. Wir denken uns diese Wege in bestimmter Reihenfolge betrachtet. W_1 ist Querschnitt von G, der zwei verschiedene Komponenten der Grenze verbindet, also ist $G - V_1$ einfach zusammenhängend; W_2 ist Querschnitt dieses Gebiets, also zerfällt $G - (V_1 + V_2)$ in zwei einfach zusammenhängende Gebiete; in einem davon liegt V_3, und W_3 zerstückelt es wieder in zwei zusammenhängende Gebiete, sodaß $G - (V_1 + V_2 + V_3)$ in drei einfach zusammenhängende Gebiete zerfällt usw. Allgemein ist (mit neuer Bedeutung von $G_1, G_2, \ldots$)

$$G - (V_1 + V_2 + \ldots + V_n) = G_1 + G_2 + \ldots + G_n$$

in n einfach zusammenhängende Gebiete zerfallen. Dabei liegt für $n \geqq 2$ das letzte V_n in einer Komponente der vorigen Zerlegung und zerstückelt sie in zwei der jetzigen, so daß bei passender Bezeichnung

$$G - (V_1 + \ldots + V_{n-1}) = G_1 + \ldots + G_{n-2} + G'_{n-1},$$
$$G'_{n-1} - V_n = G_{n-1} + G_n$$

ist. Wird die Grenze von G_i mit H_i bezeichnet, so ist also V_n mit $H_1, \ldots, H_{n-2}$ fremd, hingegen Teilmenge von H_{n-1}, H_n (nach den Formeln für H_λ und H_ϱ, S. 349) und $V_n + G_{n-1} + G_n$ ist ein Gebiet. Lassen wir diese spezielle Bezeichnung der Indices wieder fallen, so ist allgemein (da wir jedes W_i als letzten Weg betrachten können) V_i in zwei Grenzen H_k, H_l enthalten, mit den übrigen fremd,

[1] Die übrigen G_n sind, wie ähnlich folgt, einfach zusammenhängend. Die kleinen Abänderungen für den Fall, daß nicht alle G_n existieren, liegen auf der Hand.

und $\Gamma_i = V_i + G_k + G_l$ ist ein Gebiet. Umgekehrt folgt daraus, daß jede Grenze H_i zwei Mengen V_p, V_q enthalten und mit den übrigen fremd sein muß. Nämlich: H_1 kann nicht mit allen V_i fremd sein, sonst wäre

$$\mathfrak{S}(\Gamma_1, \Gamma_2, \ldots, \Gamma_n, G_2, \ldots, G_n) = V_1 + V_2 + \ldots + V_n + G_2 + \ldots + G_n$$

ein Gebiet und G damit als Summe zweier Gebiete dargestellt, also nicht zusammenhängend. H_1 kann auch nicht bloß ein einziges V_1 enthalten und mit den übrigen fremd sein, sonst wäre wieder $V_2 + \ldots + V_n + G_2 + \ldots + G_n$ ein Gebiet und $G - V_1$ nicht zusammenhängend. Also muß jedes H mindestens zwei V enthalten; enthielte aber ein H mehr als zwei V, so würde ein V in mehr als zwei H enthalten sein müssen, was dem Vorigen widerspricht. Somit ist jedes V in zwei H enthalten und jedes H enthält zwei V; danach kann man offenbar die Wege und die Gebiete zyklisch ordnen, etwa so:

$$G - (V_1 + V_2 + \ldots + V_n) = G_{12} + G_{23} + \ldots + G_{n-1\,n} + G_{n1},$$

daß die Grenze H_{ik} von G_{ik} die Mengen V_i, V_k (oder die Wege W_i, W_k) enthält; für $n = 3$ auch so:

$$G - (V_1 + V_2 + V_3) = G_1 + G_2 + G_3,$$

daß $H_i \geqq V_k$ (oder $H_i \geqq W_k$) für $i \neq k$. Die Figuren zeigen dies für den Fall, daß A, B zwei Kreisperipherien mit beschränktem Zwischengebiet sind (A im Innern von B).

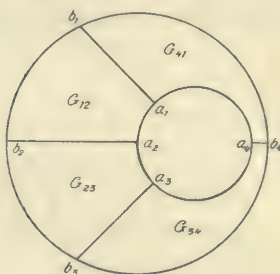


Fig. 36.

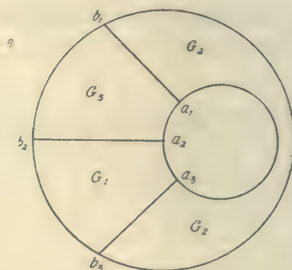


Fig. 37.

XIV. (Satz von C. Jordan.) Das umkehrbar eindeutige, stetige Bild eines Kreises ist eine geschlossene Kurve.

Um die Betrachtungen dieses Paragraphen später nicht noch einmal aufnehmen zu müssen, beweisen wir diesen Satz schon hier, obwohl wir den Begriff der stetigen Abbildung erst im folgenden Kapitel entwickeln und ihn hier durch Angabe seiner einfachsten Eigenschaften umschreiben müssen. Wir nehmen an, eine beschränkte Punktmenge C der Ebene sei auf einen Kreisumfang C_0

(an dessen Stelle auch ein Polygon treten könnte) umkehrbar eindeutig bezogen, d. h. jedem Kreispunkt a_0 entspreche ein und nur ein Punkt a von C; dabei soll (was eben die Stetigkeit ausdrückt) jeder abgeschlossenen Teilmenge von C_0 eine abgeschlossene von C und umgekehrt, jeder zusammenhängenden Teilmenge von C_0 eine zusammenhängende von C entsprechen und umgekehrt.[1] Geben wir dem Kreise einen bestimmten Umlaufsinn, so bestimmt jedes geordnete Paar a_0, b_0 verschiedener Punkte einen abgeschlossenen Kreisbogen (inkl. Endpunkte) $\overset{\frown}{a_0 b_0}$, der von $\overset{\frown}{b_0 a_0}$ verschieden ist; die entsprechende abgeschlossene Menge $\overset{\frown}{ab}$ heiße ein Bogen von C. Analog wie auf der geraden Linie erkennt man leicht, daß eine zusammenhängende Menge auf dem Kreise, wenn sie a_0, b_0 enthält, entweder den ganzen Bogen $\overset{\frown}{a_0 b_0}$ oder $\overset{\frown}{b_0 a_0}$ enthalten muß, und daß diese Forderung auch hinreicht; die einzigen abgeschlossenen zusammenhängenden Mengen $\leqq C_0$ sind, außer den einzelnen Punkten und C_0 selbst, die Kreisbögen, und die einzigen abgeschlossenen zusammenhängenden Mengen $\leqq C$, außer den einzelnen Punkten und C selbst, die Bögen. Auch ein offener Bogen ohne Endpunkte ist noch zusammenhängend; wir bezeichnen einen solchen mit $\overset{\smile}{ab}$. Da nach Tilgung von zwei verschiedenen Punkten a, b die Menge $C - \{a, b\}$ unzusammenhängend wird, so können nach S. 333 a, b keine inneren Punkte von C sein, also ist C eine Randmenge und mit der Grenze H ihres Komplementärgebiets $G = E - C$ identisch, von dem der Jordansche Satz nun behauptet, daß es in zwei Komponenten mit der gemeinsamen Grenze $C = H$ zerfällt.

Wir geben hier, in der Hauptsache, den erstaunlich einfachen Beweis von L. E. J. Brouwer wieder. Zunächst ist klar: wenn die Grenze H_1 eines zusammenhängenden Gebietes G_1 Teilmenge von H ist und ihrerseits den Bogen $\overset{\frown}{ab}$ enthält, so liegen auf diesem Bogen Punkte, die von G_1 aus geradlinig erreichbar sind. Denn ein von a, b verschiedener Punkt c des Bogens hat von dem komplementären Bogen $\overset{\frown}{ba}$ positiven Minimalabstand ϱ; nun liegen aber in jeder Umgebung von c geradlinig erreichbare Punkte, es gibt deren also auch solche, die nicht zu $\overset{\frown}{ba}$, sondern zu $\overset{\frown}{ab}$ gehören. Wenn H_1 zwei Bögen $\overset{\frown}{ab}$, $\overset{\frown}{cd}$ enthält, so läßt sich also ein Punkt des einen mit einem Punkt des andern durch einen Weg verbinden, der bis auf die Endpunkte in G_1 liegt.

(α) Jede Komponente von G hat die Grenze H.

Ist $G = E - H$ selbst zusammenhängend, so ist dies richtig. Andernfalls sei G_1 eine Komponente von G; ihre Grenze $H_1 \leqq H$

[1] Auch C selbst ist also abgeschlossen und zusammenhängend.

ist nach IX zusammenhängend und G_1 ist nach § 10, III zugleich Komponente von $E - H_1 \geqq E - H > G_1$; das Gebiet $E - H_1$ hat also noch weitere Komponenten. Eine solche sei G_2 mit der, abermals zusammenhängenden, Grenze $H_2 \leqq H_1 \leqq H$. Wenn wir also zeigen können, daß $H_2 = H$, so ist auch $H_1 = H$ und die Behauptung (α) bewiesen. Sei $H_2 < H$, so ist H_2, da es natürlich nicht aus einem einzigen Punkte bestehen kann, ein Bogen $\overset{\frown}{ad}$; wir schieben zwei Punkte b, c desselben ein und erhalten die Bögen $\overset{\frown}{ab}$, $\overset{\frown}{bc}$, $\overset{\frown}{cd}$. Die Bögen $\overset{\frown}{ab}$ und $\overset{\frown}{cd}$ lassen sich, da sie zu H_1 und H_2 gehören, durch Wege W_1 und W_2 verbinden, die bis auf die Endpunkte in G_1 und G_2 liegen. Anderseits liegen diese Wege aber auch bis auf die Endpunkte im Zwischengebiet Γ von $\overset{\frown}{ab}$, $\overset{\frown}{cd}$ und zerlegen es in zwei zusammenhängende Gebiete Γ_1, Γ_2. Jeder Punkt von V_1 ($= W_1$ minus Endpunkte) ist Häufungspunkt sowohl von Γ_1 als auch von Γ_2 und zugleich Punkt von G_1; also hat G_1 sowohl mit Γ_1 als mit Γ_2 Punkte gemein, und gleiches gilt von G_2. Da Γ_1 sowohl mit G_1 als mit G_2 Punkte gemein hat, so muß Γ_1 auch die Grenze H_2 treffen, d. h. den offenen Bogen $\underset{\smile}{bc}$, und das gleiche gilt von Γ_2. Der offene Bogen $\underset{\smile}{bc}$ trifft also die Gebiete Γ_1, Γ_2, müßte also, weil er zusammenhängend ist, auch die Grenze von Γ_1 treffen. Aber diese liegt in der Summe der Mengen $\overset{\frown}{ab}$, $\overset{\frown}{cd}$, W_1, W_2 und wird also von $\underset{\smile}{bc}$ nicht getroffen. Die Annahme $H_2 < H$ hat also zum Widerspruch geführt, es ist $H_2 = H$ und $H_1 = H$.

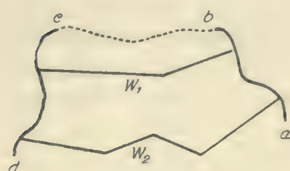


Fig. 38.

(β) G hat höchstens zwei Komponenten.

Nehmen wir an, G habe mindestens drei Komponenten G_1, G_2, G_3, jede mit der Grenze H. Wir zerlegen H in vier Bögen $\overset{\frown}{ab}$, $\overset{\frown}{bc}$, $\overset{\frown}{cd}$, $\overset{\frown}{da}$ und verbinden $\overset{\frown}{ab}$, $\overset{\frown}{cd}$ durch drei Wege W_1, W_2, W_3, die bis auf die Endpunkte in G_1, G_2, G_3 liegen. Diese Wege liegen bis auf die Endpunkte auch im Zwischengebiet Γ von $\overset{\frown}{ab}$, $\overset{\frown}{cd}$ und zerlegen es in drei zusammenhängende Gebiete Γ_1, Γ_2, Γ_3, derart, daß (bei passender Numerierung, S. 354) W_1 den Grenzen von Γ_2 und Γ_3 angehört usw. Es hat dann G_1 Punkte mit Γ_2, Γ_3 gemein, Γ_1 Punkte mit G_2, G_3 usw., jedes der Gebiete Γ_1, Γ_2, Γ_3 muß also H, d. h. einen der offenen Bögen $\underset{\smile}{bc}$, $\underset{\smile}{da}$ treffen. Mindestens einer dieser offenen Bögen, etwa $\underset{\smile}{bc}$, muß also zwei Gebiete, etwa Γ_1, Γ_2 treffen, was denselben Widerspruch wie vorhin ergibt, da der offene zu-

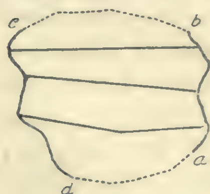


Fig. 39.

sammenhängende Bogen $\underset{\smile}{bc}$ die Mengen $\widehat{ab}$, $\widehat{cd}$, W_1, W_2, W_3 und also die Grenze von Γ_1 (oder Γ_2) nicht trifft.

(γ) G hat mindestens zwei Komponenten.

Angenommen, G sei zusammenhängend; dann ist es auch jede Menge zwischen G und $G + H = E$. Wir verbinden zwei getrennte Bögen von H durch einen bis auf die Endpunkte in G verlaufenden Weg

$$W = (a, \ldots, b) = \{a, b\} + V;$$

dies ist ein Querschnitt von G und zerlegt es in zwei zusammenhängende Gebiete

$$G - V = G_1 + G_2.$$

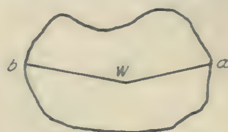


Fig. 40.

W ist aber auch ein Querschnitt in dem zusammenhängenden Gebiet $\Gamma = G + \underset{\smile}{ba} = E - \widehat{ab}$ und zerlegt es in zwei zusammenhängende Gebiete

$$\Gamma - V = \Gamma_1 + \Gamma_2,$$

woraus

$$\Gamma_1 + \Gamma_2 = G_1 + G_2 + \underset{\smile}{ba}$$

folgt. Jede der rechts stehenden zusammenhängenden Mengen gehört einer der links stehenden an, etwa

$$\Gamma_1 = G_1 + \underset{\smile}{ba}, \quad \Gamma_2 = G_2;$$

daraus folgt, daß die Punkte des offenen Bogens $\underset{\smile}{ba}$ nicht Häufungspunkte von G_2 sind; $\underset{\smile}{ba}$ ist mit H_2 fremd. Indem man drittens W als Querschnitt in $G + \underset{\smile}{ab}$ auffaßt, findet man ebenso, daß der offene Bogen $\underset{\smile}{ab}$ mit einer der Grenzen H_1, H_2 fremd ist. Der Durchschnitt $\mathfrak{D}(H_1, H_2)$ enthält also keinen Punkt von $\underset{\smile}{ab} + \underset{\smile}{ba}$. Da andererseits

$$\mathfrak{S}(H_1, H_2) = \mathfrak{S}(H, W), \quad \mathfrak{D}(H_1, H_2) \geqq W,$$

so ist $\mathfrak{D}(H_1, H_2) = W$. Dieser Durchschnitt müßte aber nach S. 347 eine geschlossene Kurve sein, da $\mathfrak{S}(H, W)$ abgeschlossene Randmenge und G_1, G_2 die Komponenten ihres Komplements sind; und das stimmt nicht, da $E - W$ zusammenhängend ist. Also ist die Voraussetzung, G sei zusammenhängend, unhaltbar, und mit (α) (β) (γ) ist der Jordansche Satz bewiesen.

Wir bemerken noch: das umkehrbar eindeutige, stetige Bild H einer Strecke[1] hat als Komplement ein zusammenhängendes

---

[1] oder eines Kreisbogens; man darf aber natürlich nicht ohne weiteres annehmen, daß ein Kreisbogenbild auch Bogen eines Kreisbildes sei, d. h. daß sich die umkehrbar eindeutige, stetige Abbildung eines Kreisbogens zu einer solchen des ganzen Kreises erweitern lasse. In diesem Fall wäre die Behauptung des Textes unmittelbare Konsequenz des Jordanschen Satzes.

Gebiet $G = E - H$ mit der Grenze H. Der Beweis nämlich, daß jede Komponente von G die Grenze H hat, verläuft wörtlich genau wie der obige Beweis von (α), und dieselbe Überlegung, in der man nunmehr unter $\widehat{ad}$ die ganze Menge H zu verstehen hat, zeigt auch noch, daß G nicht einmal zwei Komponenten haben kann.

Neuntes Kapitel.

Abbildungen oder Funktionen.

§ 1. Stetige Funktionen.

Der allgemeine Formalismus funktionaler Beziehungen zwischen Mengen ist uns aus Kap. II, § 4 bekannt; wir beschränken uns hier auf eindeutige Funktionen und deren Umkehrungen. Jedem Punkt x einer Menge A entspreche ein und nur ein Punkt $y = f(x)$, den wir das Bild von x nennen; damit ist in A eine (eindeutige) Funktion $f(x)$ definiert. Die Menge dieser Bilder sei B; jedem Punkt y von B entsprechen ein oder mehrere Punkte $x = \varphi(y)$ von A, deren Bild y ist und die wir Urbilder von y nennen; die inverse Funktion $\varphi(y)$ ist also im allgemeinen mehrdeutig. Ist auch $\varphi(y)$ eindeutig, d. h. hat jeder Punkt y ein einziges Urbild, so heißt jede der beiden Funktionen umkehrbar eindeutig (eineindeutig). Ist P eine beliebige Menge $\leqq A$, so bezeichnen wir die Menge $F(P)$ der Bilder aller Punkte von P als das Bild[1] von P; ebenso sei, für $Q \leqq B$, $\Phi(Q)$ die Menge aller Urbilder aller Punkte von Q, das Urbild von Q. Wir erinnern an die damaligen Formeln:

(1) $$\Phi(F(P)) \geqq P, \quad F(\Phi(Q)) = Q;$$

(2) $$\begin{cases} F(\mathfrak{S}(P_1, P_2, \ldots)) = \mathfrak{S}(F(P_1), F(P_2), \ldots), \\ F(\mathfrak{D}(P_1, P_2, \ldots)) \leqq \mathfrak{D}(F(P_1), F(P_2), \ldots), \\ \Phi(\mathfrak{S}(Q_1, Q_2, \ldots)) = \mathfrak{S}(\Phi(Q_1), \Phi(Q_2), \ldots), \\ \Phi(\mathfrak{D}(Q_1, Q_2, \ldots)) = \mathfrak{D}(\Phi(Q_1), \Phi(Q_2), \ldots). \end{cases}$$

Bisweilen ist die ausdrückliche Angabe, daß x die Menge A durchlaufen soll, erwünscht; wir schreiben dann $f(x|A)$ statt $f(x)$. Läßt man x nur eine Teilmenge A' von A durchlaufen, so ist hiermit eine Teilfunktion $f(x|A')$ aus der Gesamtfunktion ausgesondert.

---

[1] Nötigenfalls können die Bilder von Punkten und Mengen als Bildpunkte und Bildmengen unterschieden werden.

Bekanntlich nennt man eine reelle Funktion $f(x)$ einer reellen Variablen an der Stelle a **stetig**, wenn sich zu jedem vorgeschriebenen $\sigma > 0$ ein $\varrho > 0$ derart wählen läßt, daß aus $|x - a| < \varrho$ stets $|f(x) - f(a)| < \sigma$ folgt. Diese beiden Ungleichungen definieren die **Umgebung** U_a des Punktes a mit dem Radius ϱ und einen Teil der Umgebung V_b des Punktes $b = f(a)$ mit dem Radius σ; die Stetigkeitsforderung besagt also, daß zu gegebenem V_b ein U_a so gewählt werden kann, daß die Bilder aller Punkte von U_a in V_b liegen. Dies nehmen wir ebenfalls zur allgemeinen Definition der Stetigkeit, wobei zunächst also A, B nur als topologische Räume vorausgesetzt zu werden brauchen, in denen die Umgebungsaxiome gelten.

Definition. Die Funktion $y = f(x)$ heißt im Punkte a stetig, wenn zu jeder Umgebung V_b des Punktes $b = f(a)$ eine Umgebung U_a des Punktes a existiert, deren Bild in V_b liegt: $F(U_a) \leqq V_b$.

Eine prinzipielle Bemerkung ist hier vorauszuschicken: wenn wir A, B selbst als die von x, y durchlaufenen Räume ansehen, so kümmern wir uns nicht um etwaige sonst vorhandene Punkte; U_a bedeutet eine Teilmenge von A, V_b eine von B. Andererseits kann es, z. B. bei Betrachtung mehrerer Funktionen, notwendig werden, A als Teilmenge eines Raumes E_x, B als Teilmenge desselben oder eines andern Raumes E_y aufzufassen, womit unsere gegenwärtige Betrachtung zu einer **Relativtheorie** bezüglich A und B wird (Kap. VII, § 6); unter U_a ist also dann eine Relativumgebung, d. h. der Durchschnitt einer absoluten (in E_x liegenden) Umgebung mit A zu verstehen, unter einer abgeschlossenen Teilmenge von A eine in A abgeschlossene Menge usw. Um unsere Sätze so zu formulieren, daß sie in dieser Hinsicht keinem Mißverständnis ausgesetzt sind, wollen wir an Stellen, wo es wünschenswert erscheint, den entsprechenden Relativitätsvermerk (in A abgeschlossen usw.) beifügen.

In einem **isolierten** Punkt a von A ist jede Funktion stetig, denn es gibt hier eine Umgebung $U_a = \{a\}$, die aus a allein besteht, und deren Bild $\{b\}$ in jedem V_b liegt. Nur ein Häufungspunkt von A legt also einer daselbst stetigen Funktion eine Beschränkung auf.

Ist A' eine den Punkt a enthaltende Teilmenge von A und ist die Gesamtfunktion $f(x|A)$ in a stetig, so ist auch die Teilfunktion $f(x|A')$ in a stetig.

Ist $z = g(y)$ eine in $B' \geqq B$ definierte Funktion, so ist

$$z = h(x) = g\left(f(x)\right)$$

eine in A definierte Funktion; man kann hierbei offenbar $B' = B$ annehmen, da nur die Teilfunktion $g(y|B)$ in Betracht kommt. Ist

$f(x)$ in a und $g(y)$ in $b = f(a)$ stetig, so ist die zusammengesetzte Funktion $h(x)$ in a stetig; denn einer beliebigen Umgebung W_c entspricht eine Umgebung V_b und dieser eine Umgebung U_a derart, daß

$$G(V_b) \leqq W_c, \quad F(U_a) \leqq V_b, \quad H(U_a) = G(F(U_a)) \leqq W_c.$$

Kurz ausgedrückt: eine stetige Funktion einer stetigen Funktion ist wieder eine stetige Funktion.

Die Stetigkeitsforderung $F(U_a) \leqq V_b$ ist nach (1) mit $U_a \leqq \Phi(V_b)$ gleichbedeutend; denn aus der ersten Formel folgt

$$U_a \leqq \Phi(F(U_a)) \leqq \Phi(V_b)$$

und aus der zweiten $F(U_a) \leqq F(\Phi(V_b)) = V_b$. Sie besagt in dieser neuen Form, daß das Urbild jeder Umgebung V_b den Punkt a zum inneren Punkt haben soll, oder:

I. Die Funktion $f(x)$ ist in a stetig dann und nur dann, wenn jeder Teilmenge Q von B, die den Punkt $b = f(a)$ zum inneren Punkt hat, eine Urbildmenge $\Phi(Q)$ entspricht, die den Punkt a zum inneren Punkt hat.

Umgekehrt läßt sich dies wieder auf folgende Form bringen:

II. Die Funktion $f(x)$ ist in a stetig dann und nur dann, wenn jeder Teilmenge P von A, die den Punkt a zum α-Punkt hat, eine Bildmenge $F(P)$ entspricht, die den Punkt $b = f(a)$ zum α-Punkt hat.

Um dies zu beweisen, ist die Gleichwertigkeit der Q-Bedingung in I und der P-Bedingung in II zu zeigen.

Ist die Q-Bedingung erfüllt, so sei P eine Menge, die a zum α-Punkt hat. Wir setzen $Q = F(P)$; nach (2) ist

$$A = \Phi(B) = \Phi(Q) + \Phi(B - Q)$$

und nach (1) ist $\Phi(Q) \geqq P$, also $\Phi(B - Q) \leqq A - P$. Wäre nun b kein α-Punkt von Q, so wäre er innerer Punkt von $B - Q$, also a innerer Punkt von $\Phi(B - Q)$ oder von $A - P$ und demgemäß kein α-Punkt von P. Also ist b ein α-Punkt von Q; die P-Bedingung folgt aus der Q-Bedingung.

Ist die P-Bedingung erfüllt, so sei Q eine Menge, die b zum inneren Punkt hat. Wir setzen $P = \Phi(Q)$, also $A - P = \Phi(B - Q)$ und nach (1) wieder $F(A - P) = B - Q$. Wäre nun a kein innerer Punkt von P, so wäre er α-Punkt von $A - P$, also b wäre α-Punkt von $B - Q$ und kein innerer Punkt von Q. Also ist a innerer Punkt von P; die Q-Bedingung folgt aus der P-Bedingung.

Man nennt die Funktion $f(x)$ schlechthin stetig (genauer: in A stetig), wenn sie in allen Punkten von A stetig ist; B heiße dann ein stetiges Bild von A. Dazu ist also notwendig und hinreichend:

nach I, daß die Urbilder der inneren Punkte von Q innere Punkte des Urbilds $\Phi(Q)$ seien, für jede Teilmenge Q von B, also

(3) $$\Phi(Q_i) \leqq \Phi(Q)_i;$$

nach II, daß die Bilder der α-Punkte von P α-Punkte des Bildes $F(P)$ seien, für jede Teilmenge P von A, also

(4) $$F(P_\alpha) \leqq F(P)_\alpha.$$

Dies gestattet folgende bequemere Formulierung:

III. Die Funktion $f(x)$ ist dann und nur dann stetig, wenn jedem Relativgebiet von B als Urbild ein Relativgebiet von A, oder jeder in B abgeschlossenen Menge als Urbild eine in A abgeschlossene Menge entspricht.

Ist nämlich $f(x)$ stetig und $Q = Q_i$ ein Gebiet, so ist $\Phi(Q) \leqq \Phi(Q)_i$ oder $\Phi(Q) = \Phi(Q)_i$, also auch $\Phi(Q)$ ein Gebiet. Wenn umgekehrt jedem Gebiet Q ein Gebiet $\Phi(Q)$ entspricht, so ist, für eine beliebige Menge Q, $\Phi(Q_i)$ ein in $\Phi(Q)$ enthaltenes Gebiet, also $\Phi(Q_i) \leqq \Phi(Q)_i$, $f(x)$ stetig. Die zweite Form des Satzes folgt durch Komplementbildung wegen $\Phi(B - Q) = A - \Phi(Q)$, oder aus der Formel (4). Man beachte wohl: die Bedingung lautet nicht so, daß einer in A abgeschlossenen Menge als Bild eine in B abgeschlossene entsprechen soll, sondern umgekehrt.

Wir geben von dem Satze III eine Anwendung auf reelle Funktionen (vgl. auch § 5). $f(x)$ sei eine reelle, in A stetige Funktion, b eine reelle Zahl und L, M, N seien die Mengen derjenigen Punkte x von A, wo $f(x)$ kleiner, gleich, größer als b ist $(A = L + M + N)$. Dann folgt aus III, daß die Mengen M, $L + M$, $M + N$ in A abgeschlossen sind; denn es ist z. B. $L + M$ das Urbild der Menge derjenigen zu $B = F(A)$ gehörigen reellen Zahlen, die $\leqq b$ sind, und diese Menge ist in B abgeschlossen. Nehmen wir A als den Raum der Punkte x und unterdrücken den Relativitätsvermerk, so sind die obengenannten Mengen (absolut) abgeschlossen, ihre Komplemente $L, N, L + N$ sind Gebiete.[1] Ein Punkt a von M, wo also $f(a) = b$, heißt eine Konstanzstelle der Funktion, wenn eine Umgebung U_a existiert, wo durchweg $f(x) = b$, also wenn a innerer Punkt von M ist; die Menge der zu diesem Funktionswerte gehörigen Konstanzstellen ist das Gebiet M_i. Ferner heißt a eine Maximalstelle, wenn eine Umgebung U_a vorhanden ist, in der $f(x) \leqq b$, aber nicht durchweg $f(x) = b$; insbesondere eine eigentliche Maximalstelle, wenn eine Umgebung U_a existiert, in der, für $x \neq a$, $f(x) < b$. Entsprechendes gilt für Minima. Um die

[1] Abgesehen von diesem durch die Stetigkeit von $f(x)$ bedingten Charakter der auftretenden Mengen gilt das Folgende auch für unstetige Funktionen.

Mengen dieser (zum Funktionswert b gehörigen) Punkte darzustellen, setzen wir

$$L_1 = \mathfrak{D}(M, L_\beta), \quad M_1 = \mathfrak{D}(M, M_\beta), \quad N_1 = \mathfrak{D}(M, N_\beta),$$
$$M = L_1 + L_2 = M_1 + M_2 = N_1 + N_2$$
$$= L_0 \qquad = M_0 \qquad = N_0$$

und

$$M_{\lambda\mu\nu} = \mathfrak{D}(L_\lambda, M_\mu, N_\nu) \qquad (\lambda, \mu, \nu = 0, 1, 2).$$

Dann ist:

M_{122} die Menge der eigentlichen Maximalstellen,
M_{112} „ „ „ uneigentlichen „
M_{102} „ „ „ sämtlichen „
M_{221} „ „ „ eigentlichen Minimalstellen,
M_{211} „ „ „ uneigentlichen „
M_{201} „ „ „ sämtlichen „
M_{202} „ „ „ Konstanzstellen[1],
M_{101} „ „ „ gewöhnlichen Stellen,

d. h. derer, die weder Maximal- noch Minimal- noch Konstanzstellen sind. In der Tat sind die eigentlichen Maximalstellen diejenigen Punkte von M, die Häufungspunkte von L, aber nicht von M und N sind (die also gleichzeitig zu L_1, M_2, N_2 gehören); die uneigentlichen Maximalstellen sind Häufungspunkte von L und M, aber nicht von N; die Maximalstellen überhaupt Häufungspunkte von L, aber nicht von N; die Konstanzstellen Häufungspunkte weder von L noch von N, die gewöhnlichen Stellen Häufungspunkte sowohl von L als auch von N. (Zu den Konstanzstellen werden natürlich auch etwaige in M liegende isolierte Punkte von A gerechnet.) Alle diese Mengen sind Differenzen abgeschlossener Mengen; denn die Mengen mit dem Index 1 sind abgeschlossen (in M oder absolut), die mit dem Index 2 Relativgebiete von M; also ist, wenn wie immer F abgeschlossene Mengen und G Gebiete bedeutet, $M_{\lambda\mu\nu}$ entweder ein F oder von der Form $\mathfrak{D}(F, G) = F - F'$. Auch die Menge aller Extremalstellen

$$M_{102} + M_{201} = \mathfrak{D}(L_1, N_2) + \mathfrak{D}(L_2, N_1) = \mathfrak{S}(L_1, N_1) - \mathfrak{D}(L_1, N_1)$$

und noch andere Mengen, die man durch Zusammenfassung der obengenannten bilden kann, sind Differenzen abgeschlossener Mengen.

Dies alles bezog sich auf einen einzigen Funktionswert b, und die definierten Mengen können natürlich auch teilweise oder sämtlich (wenn $M = 0$) verschwinden. Nur wenn es wirklich eine Maximalstelle a mit $f(a) = b$ gibt, also M_{102} nicht verschwindet, nennen wir b einen Maximalwert oder ein Maximum der Funktion, und entsprechend einen Konstanzwert, einen eigentlichen Maximalwert usw.

[1] Daß dies mit M_i übereinstimmt, ist leicht zu sehen.

Nun ist leicht zu sehen, daß die Menge der Maximalwerte der Funktion **höchstens abzählbar** ist, falls im Raume A das zweite Abzählbarkeitsaxiom gilt. Denn sind b_1 und b_2 zwei verschiedene Maximalwerte, a_1 und a_2 zwei zugehörige Maximalstellen und U_{a_1}, U_{a_2} solche Umgebungen von diesen, daß in ihnen resp. $f(x) \leqq b_1$ und $f(x) \leqq b_2$, so müssen diese Umgebungen verschieden sein, da andernfalls a_1 zu U_{a_2} gehören, also $b_1 = f(a_1) \leqq b_2$ und ebenso gleichzeitig $b_2 \leqq b_1$, also $b_1 = b_2$ sein müßte. Es läßt sich also jedem Maximalwert b eine Umgebung U_a in A so zuordnen, daß verschiedenen Maximalwerten verschiedene Umgebungen entsprechen, und die Menge der b ist höchstens abzählbar. Die **Menge aller Maximalstellen** von $f(x)$ ist demnach eine Summe von höchstens abzählbar vielen Differenzen abgeschlossener Mengen (in einem metrischen Raume ein F_σ), und das gleiche gilt von den eigentlichen Maximalstellen, Konstanzstellen usw. Die Menge der gewöhnlichen Stellen ist, wenn A metrisch ist, ein G_δ.

Kehren wir nach diesem Exkurse zur allgemeinen Theorie zurück.

IV. **Das stetige Bild einer zusammenhängenden Menge ist wieder zusammenhängend. Das stetige Bild einer beliebigen Menge hat nicht mehr Komponenten und Quasikomponenten als diese.**

Ist $f(x)$ stetig, so muß einer Zerlegung $B = Q + Q'$ in zwei relativ abgeschlossene, nichtverschwindende Teilmengen eine ebensolche $A = \Phi(Q) + \Phi(Q') = P + P'$ entsprechen; ist also A zusammenhängend, d. h. keine solche Zerlegung möglich, so ist auch B zusammenhängend. Da $f(x)$ eine stetige Funktion bleibt, wenn man x nur eine Teilmenge von A durchlaufen läßt, so entspricht jeder zusammenhängenden Menge $\leqq A$ eine ebensolche $\leqq B$; einer Komponente P von A entspricht ein Bild $Q = F(P)$, das ganz in eine Komponente von B fällt, während natürlich recht wohl mehrere solche Bilder derselben Komponente von B angehören können; die Komponentenzahl von B ist also höchstens[1] gleich der von A. Gehören zwei Punkte q, q' von B verschiedenen Quasikomponenten an, sodaß, bei einer der obigen Zerlegungen $B = Q + Q'$, q zu Q und q' zu Q' gehört, so gehört jedes Urbild $p = \varphi(q)$ zu P und jedes $p' = \varphi(q')$ zu P'; p und p' gehören zu verschiedenen Quasikomponenten von A. Das Bild einer Quasikomponente von A fällt also in eine Quasikomponente von B; B hat nicht mehr Quasikomponenten als A.

---

[1] Sie kann natürlich kleiner und insbesondere das stetige Bild einer unzusammenhängenden Menge auch zusammenhängend sein; z. B. wenn $f(x)$ konstant ist, B sich auf einen Punkt reduziert.

Ein spezieller Fall von IV ist der bekannte Satz, daß eine, in einer zusammenhängenden Menge A (z. B. in einem reellen Zahlenintervall) stetige reelle Funktion $f(x)$ jeden Wert zwischen zweien ihrer Werte annimmt; denn die zusammenhängende Menge B reeller Zahlen hat diese Eigenschaft (S. 259).

Ist $f(x)$ umkehrbar eindeutig und beiderseits stetig (d. h. $f(x)$ und $\varphi(y)$ stetig), so entspricht jeder Komponente von A eine Komponente von B und umgekehrt, ebenso jeder Quasikomponente; beide Mengen haben die gleiche Komponenten- und Quasikomponentenzahl.

V. Das stetige Bild einer in sich kompakten Menge ist wieder in sich kompakt.

A sei kompakt (im Raume A, d. h. in sich; jede unendliche Teilmenge von A hat einen Häufungspunkt in A). Ist Q eine unendliche Teilmenge von B, so betrachten wir von jedem ihrer Punkte q ein einziges Urbild $p = \varphi(q)$; diese bilden die eineindeutig auf Q bezogene Menge P, wobei $Q = F(P)$. Nun hat P einen Häufungspunkt a; dieser ist α-Punkt von P und jeder durch Weglassung endlich vieler Punkte daraus entstehenden Teilmenge P'. Nach II hat $b = f(a)$ die gleiche Eigenschaft gegenüber Q, ist also Häufungspunkt von Q. (Oder, auf Grund der ursprünglichen Stetigkeitsdefinition: ist V_b beliebig und $F(U_a) \leqq V_b$, so liegen in U_a unendlich viele Punkte von P, also in V_b unendlich viele Punkte von Q.) B ist also in sich kompakt.

Ist z. B. $f(x)$ eine reelle stetige Funktion, A eine in sich kompakte (d. h. abgeschlossene und beschränkte) Menge im euklidischen Raum, so ist B eine in sich kompakte (d. h. abgeschlossene und beschränkte) Menge reeller Zahlen, hat also ein Maximum und ein Minimum (Weierstrass).

Nehmen wir jetzt für die Räume A, B das erste Abzählbarkeitsaxiom (Kap. VIII, § 2) hinzu, so werden die α-Punkte einer Menge mit den Limites konvergenter Punktfolgen aus dieser Menge identisch, und wir können den bisweilen als Definition der Stetigkeit verwendeten Satz aussprechen:

VI. Die Funktion $f(x)$ ist in a stetig dann und nur dann, wenn jeder nach a konvergenten Folge von Punkten a_n eine nach $b = f(a)$ konvergente Folge von Bildpunkten $b_n = f(a_n)$ entspricht.

Denn mit dieser Limesbedingung ist auch die in II erfüllt; ist andererseits $f(x)$ in a stetig und $\lim a_n = a$, so ist a ein α-Punkt der Folge a_n und jeder Teilfolge, also b ebenfalls α-Punkt der Folge b_n und jeder Teilfolge, d. h. $b = \lim b_n$ (oder: ist V_b eine

beliebige Umgebung und $F(U_a) \leqq V_b$, so liegen fast alle a_n in U_a, also fast alle b_n in b).

VII. Eine in A stetige Funktion $f(x)$ ist bekannt, wenn man die Bilder der Punkte einer in A dichten Teilmenge P kennt.

Denn wegen $A = P_\alpha$ kann man jeden Punkt von A als Limes einer konvergenten Folge von Punkten aus P darstellen ($a = \lim p_n$); dann ist $f(a) = \lim f(p_n)$ eindeutig bestimmt.

Z. B. ist eine stetige Funktion einer reellen Variablen durch ihre Werte an den rationalen Stellen bestimmt.

Haben A und P die Mächtigkeiten $\mathfrak{a}$, $\mathfrak{p}$, und werden alle in A definierten Funktionen $f(x)$ in Betracht gezogen, bei denen $f(x)$ Element einer Menge C von der Mächtigkeit $\mathfrak{c}$ ist, so hat die Gesamtheit dieser Funktionen die Mächtigkeit $\mathfrak{c}^{\mathfrak{a}}$, die der stetigen Funktionen aber nur eine Mächtigkeit $\leqq \mathfrak{c}^{\mathfrak{p}}$ (nicht jede in P definierte Funktion $f(x)$ braucht ja zu einer in A stetigen Funktion erweiterungsfähig zu sein; eine notwendige Bedingung dafür ist jedenfalls, daß $f(x)$ in P stetig sei). In einem Raum A mit abzählbarer dichter Teilmenge hat die Menge aller in A stetigen Funktionen also höchstens die Mächtigkeit $\mathfrak{c}^{\aleph_0}$. So hat die Menge aller reellen stetigen Funktionen einer reellen Variablen nur die Mächtigkeit $\aleph^{\aleph_0} = \aleph$ ($\aleph$ Mächtigkeit des Kontinuums), die aller Funktionen (selbst solcher, die nur zwei Werte annehmen können) die Mächtigkeit $\aleph^{\aleph} = 2^{\aleph} > \aleph$. Stetigkeit ist für Funktionen also eine analoge Beschränkung wie Abgeschlossenheit für Mengen.

VIII. Ist B umkehrbar eindeutiges, stetiges Bild der in sich kompakten Menge A, so ist auch A stetiges Bild von B.

Denn sind $y = f(x)$ und $x = \varphi(y)$ eindeutige Funktionen und A kompakt, so entspricht jeder abgeschlossenen, folglich in sich kompakten Teilmenge P von A nach V eine in sich kompakte, folglich abgeschlossene Teilmenge Q von B (vgl. S. 264), und nach III ist daher die Funktion $\varphi(y)$ stetig.

Sind A, B metrische Räume, wie wir von jetzt an voraussetzen, so ist die Bedingung für Stetigkeit in a diese, daß zugleich mit der Entfernung $\overline{ax_n}$ die Entfernung $\overline{by_n}$ der Bilder nach Null konvergiert. Ist dies für jedes a der Fall, so ist die Funktion in A stetig; ist es nicht nur bei festgehaltenem, sondern auch bei variablem a der Fall, d. h. konvergiert mit $\overline{a_n x_n}$ auch $\overline{b_n y_n}$ nach Null, so ist $f(x)$ gleichmäßig stetig (s. unten). Dies trifft ein, wenn man eine Relation der Form $\overline{by} \leqq \sigma(\overline{ax})$ beweisen kann, wo $\sigma(\xi)$ eine

positive, mit der positiven Variablen ξ nach 0 konvergierende Funktion ist, z. B. $\sigma(\xi) = \xi,\ 2\xi,\ \sqrt{\xi}$ u. dgl.

Beispiele. Denken wir uns A, B als Teilmengen metrischer Räume E_x, E_y. Die untere Entfernung $y = \delta(x, M)$ eines Punktes x in A von einer beliebigen nichtverschwindenden Menge $M \leqq E_x$ ist eine stetige Funktion von x, denn nach § 6, (4) ist

$$\overline{yy'} = |y' - y| = |\delta(x', M) - \delta(x, M)| \leqq \overline{xx'}.$$

Verschiedene frühere Sätze, die sich auf den Fall einer abgeschlossenen kompakten Menge A beziehen, erscheinen hiernach als Spezialfälle von V.

Im euklidischen Raume gibt die senkrechte Projektion einer Menge A auf eine Gerade, Ebene usw. ein stetiges Bild von A, da auch hier $yy' \leqq xx'$ ist. Jede rechtwinklige Koordinate des Punktes x ist eine stetige Funktion von x.

Ist $y = f(x)$ stetig und E_y ein n-dimensionaler euklidischer Raum, so werden die Koordinaten von y stetige Funktionen von y, also von x; $y = f(x)$ repräsentiert demnach n reelle stetige Funktionen von x

$$y_1 = f_1(x),\ y_2 = f_2(x),\ \ldots,\ y_n = f_n(x).$$

Ist $y = f(x)$ eine reelle stetige Funktion und E_x ein m-dimensionaler euklidischer Raum, so fassen wir $f(x)$ als Funktion des reellen Zahlenkomplexes $x = (x_1, x_2, \ldots, x_m)$ auf und nennen

$$y = f(x_1, x_2, \ldots, x_m)$$

eine in A definierte stetige Funktion der reellen Variablen $x_1, x_2, \ldots, x_m$. Geben wir den Variablen $x_2, \ldots, x_m$ konstante Werte $a_2, \ldots, a_m$, d. h. betrachten eine der ersten Koordinatenachse parallele Gerade T, so ist, falls auf dieser Punkte von A liegen, $f(x_1, a_2, \ldots, a_m)$ eine in $A' = \mathfrak{D}(A, T)$ definierte stetige Funktion von x_1; die Funktion ist in jeder einzelnen Variablen partiell stetig. Umgekehrt läßt sich aus der Stetigkeit in jeder einzelnen Variablen die totale Stetigkeit nicht erschließen; denn daraus, daß bei geradliniger, den Achsen paralleler Konvergenz von x nach a immer $f(x)$ nach $f(a)$ konvergiert, folgt natürlich keineswegs, daß dies auch bei beliebiger Konvergenz von x nach a der Fall sein wird.

Eine stetige Funktion $y = f(x)$, wo E_x m-dimensional und E_y n-dimensional ist, repräsentiert n reelle stetige Funktionen von m reellen Variablen.

Ist E_x ein Hilbertscher Raum, so erhalten wir stetige Funktionen von abzählbar unendlich vielen reellen Variablen $x_1, x_2, \ldots$. Aber auch der Fall, daß die Elemente von E_x Punktmengen oder Funktionen sind, subsumiert sich unter unsere allgemeine Betrachtung; sind z. B. die $x = x(t)$ reelle, im Intervall $\alpha \leqq t \leqq \beta$ definierte, stetige Funktionen der

reellen Variablen t und definiert man (S. 289) die Entfernung $\overline{xx'}$ als Maximum von $|x(t) - x'(t)|$, so ist

$$y = f(x) = \int_\alpha^\beta x(t)\,dt$$

eine stetige Funktion von x. Denn man hat

$$|y - y'| \leqq \int_\alpha^\beta |x(t) - x'(t)|\,dt \leqq \overline{xx'}\,.\,(\beta - \alpha),$$

mit $\overline{xx'}$ konvergiert also auch $\overline{yy'}$ nach Null.

Gleichmäßige Stetigkeit. Sind A, B metrische Räume und $y = f(x)$ eine stetige Funktion, so ermöglicht die Vergleichbarkeit der Umgebungen verschiedener Punkte (nach dem Radius) eine Vergleichung der Stetigkeit in verschiedenen Punkten. Ist σ eine vorgegebene positive Zahl, so gibt es für jedes x eine Umgebung U_x, vom Radius ϱ_x, deren Bild in die Umgebung V_y vom Radius σ fällt. Dieser Radius ϱ_x gibt ein Maß für den Grad der Stetigkeit, wenn man so sagen will, im Punkte x: die Stetigkeit in x ist um so größer oder geringer, je größer man, bei vorgeschriebenem σ, ϱ_x wählen kann oder je kleiner man es wählen muß. Reicht man für alle Punkte x mit einem und demselben $\varrho_x = \varrho$ aus, und gilt dies für jedes σ, so heißt $f(x)$ in A gleichmäßig stetig; auf jedes vorgegebene $\sigma > 0$ läßt sich also dann ein $\varrho > 0$ derart finden, daß aus $\overline{xx'} < \varrho$ stets $\overline{yy'} < \sigma$ folgt, gleichviel welche Punkte x, x' von A man wählen möge. Es gilt der zu VI analoge Satz:

IX. Die Funktion $f(x)$ ist gleichmäßig stetig dann und nur dann, wenn jeder Folge von Punktpaaren x_n, x_n', deren Entfernung nach Null konvergiert, eine Folge von Bildpaaren y_n, y_n' entspricht, deren Entfernung nach Null konvergiert.

Es braucht nur bewiesen zu werden, daß diese Bedingung hinreichend ist. Ist $f(x)$ nicht gleichmäßig stetig, so gibt es zu einem gewissen σ kein solches ϱ, wie oben gefordert, d. h. für dieses σ und jedes ϱ gibt es mindestens ein Punktpaar $\overline{xx'} < \varrho$, für dessen Bild $\overline{yy'} \geqq \sigma$. Es gibt also für jede natürliche Zahl n ein Punktpaar $\overline{x_n x_n'} < \frac{1}{n}$, $\overline{y_n y_n'} \geqq \sigma$, und dann ist zwar $\lim \overline{x_n x_n'} = 0$, hingegen nicht $\lim \overline{y_n y_n'} = 0$.

X. Ist A in sich kompakt, so ist eine in A stetige Funktion gleichmäßig stetig.

Nehmen wir eine Folge von Punktpaaren mit $\lim \overline{x_n x_n'} = 0$.

Die Folge der x_n hat einen zu A gehörigen Häufungspunkt x; für eine geeignete Teilfolge natürlicher Zahlen p ist $\lim x_p = x$, also auch $\lim x_p' = x$. Für die Bilder gilt dann $\lim y_p = \lim y_p' = y$, $\lim \overline{y_p y_p'} = 0$. Die Folge der $\overline{y_n y_n'}$ hat also eine nach Null konvergente Teilfolge, und da von jeder Teilfolge dasselbe gilt, so ist $\lim \overline{y_n y_n'} = 0$; nach IX ist $f(x)$ in A gleichmäßig stetig.

Z. B. ist eine in einem abgeschlossenen Zahlenintervall $(\alpha \leqq x \leqq \beta)$ stetige Funktion gleichmäßig stetig, was bekanntlich beim Existenzbeweise des bestimmten Integrals eine Rolle spielt. Die Funktion $f(x) = x^2$ ist in der unbeschränkten Menge der reellen Zahlen, die Funktion $f(x) = \frac{1}{x}$ in dem nicht abgeschlossenen Intervall $0 < x \leqq 1$ stetig, aber nicht gleichmäßig.

Wir sahen (VII), daß eine in A stetige Funktion $f(a)$ durch ihre Werte $f(p)$ in den Punkten einer in A dichten Teilmenge $P (P_\alpha = A)$ bestimmt ist; fragen wir umgekehrt, wann eine in P gegebene Funktion $f(p)$ zu einer in $A = P_\alpha$ stetigen Funktion $f(a)$ erweitert werden kann. Dazu ist notwendig, daß $f(p)$ in P stetig sei; hinreichend, wie wir jetzt zeigen wollen, daß $f(p)$ in P gleichmäßig stetig sei; überdies ist auch diese zweite Bedingung notwendig, falls A in sich kompakt ist, weil dann nach X die gesuchte Funktion in A und natürlich in jeder Teilmenge gleichmäßig stetig ist. Um zu zeigen, daß eine in P gleichmäßig stetige Funktion zu einer in A stetigen erweitert werden kann, müssen wir den Raum E_y, in dem die Bildpunkte $q = f(p)$ liegen, als vollständig voraussetzen resp. zuvor zu einem solchen erweitern (S. 315). Ist $a = \lim p_n$ ein beliebiger Punkt von A, so gilt für eine beliebige Folge wachsender natürlicher Zahlen $\lambda_1 < \lambda_2 < \ldots$: $\lim \overline{p_n p_{\lambda_n}} = 0$, also für die Bilder wegen der gleichmäßigen Stetigkeit von $f(x)$ in P: $\lim \overline{q_n q_{\lambda_n}} = 0$. Die Folge q_n ist also eine Fundamentalfolge und hat nach dem Cauchyschen Theorem im vollständigen Raume E_y einen Limes $b = \lim q_n$; diesen definieren wir als Bild von a: $b = f(a)$. Daß damit nur ein einziges $f(a)$ definiert wird, folgt wieder aus der gleichmäßigen Stetigkeit: ist gleichzeitig $a = \lim p_n'$, so ist $\lim \overline{p_n p_n'} = 0$, also $\lim \overline{q_n q_n'} = 0$, also $\lim q_n' = \lim q_n$. Sollte $a = p$ selbst der Menge P angehören, so ist $\lim q_n = q = f(p)$, die neue Definition von $f(a)$ stimmt also mit der alten überein. Damit ist also eine in A eindeutige Funktion $f(a)$ definiert, von der $f(p)$ eine Teilfunktion ist. Es bleibt die Stetigkeit von $f(a)$ zu zeigen, d. h. daß für $\lim a_m = a$ auch $\lim b_m = b$ ist. Sei $a_m = \lim\limits_n p_{mn}$, $b_m = \lim\limits_n q_{mn}$. Man kann dann, für jedes m, einen (von m abhängigen) Index n so angeben, daß der Punkt $p_{mn} = p_m$

und sein Bild $q_{mn} = q_m$ gleichzeitig den Bedingungen $\overline{a_m p_m} < \frac{1}{m}$, $\overline{b_m q_m} < \frac{1}{m}$ genügen. Dann ist $\lim \overline{a_m p_m} = \lim \overline{b_m q_m} = 0$, also

$$a = \lim a_m = \lim p_m,$$

demzufolge $b = \lim q_m = \lim b_m$.

§ 2. Kurven. Dimensionenzahl.

Wir setzen A, B als Punktmengen in euklidischen Räumen E_x, E_y voraus. Betrachten wir zunächst das stetige Bild einer Strecke A, die als das reelle Zahlenintervall $0 \leqq x \leqq 1$ angenommen werden kann; $y = f(x)$ sei die stetige Funktion, die bei Spaltung in Koordinaten n reelle in A stetige Funktionen

(1) $$y_1 = f_1(x), \ldots, y_n = f_n(x)$$

repräsentiert. Das Streckenbild B ist nach den Sätzen von § 1 eine beschränkte, abgeschlossene, zusammenhängende Menge, die also, falls sie nicht aus einem einzigen Punkt besteht, perfekt ist.

Man pflegt eine solche Menge als stetige Kurve zu bezeichnen.[1] Aber dieser Begriff umfaßt eine Fülle der Gestalten, die sich von dem elementargeometrischen Bilde einer „Kurve" himmelweit entfernen. Eine stetige Kurve kann Flächenstücke enthalten: das ist eine der merkwürdigsten Tatsachen der Mengenlehre, deren Entdeckung wir G. Peano verdanken. Wir zeigen (der von D. Hilbert gegebenen geometrischen Darstellung folgend), daß das eindeutige stetige Bild einer Strecke z. B. ein ganzes Quadrat sein kann.

Es sei B ein Quadrat von der Seitenlänge 1 (hier und im folgenden ist unter einem Quadrat die abgeschlossene Fläche, Inneres und Rand, unter einer Strecke eine abgeschlossene Strecke zu verstehen). Durch Halbierung der Seiten erhalten wir vier Teilquadrate, die wir so numerieren, daß in der Anordnung $B_1\ B_2\ B_3\ B_4$ nebeneinanderstehende Quadrate benachbart sind, d. h. eine Seite gemein haben. Jedes B_i wird wieder durch Seitenhalbierung in vier Quadrate B_{ik} geteilt so, daß bei der lexikographischen Anordnung

$$B_{11}\ B_{12}\ B_{13}\ B_{14}\quad B_{21}\ B_{22}\ B_{23}\ B_{24}\quad B_{31}\ B_{32}\ B_{33}\ B_{34}\quad B_{41}\ B_{42}\ B_{43}\ B_{44}$$

nebeneinanderstehende Quadrate benachbart sind; jedes B_{ik} wieder

[1] Wir geben keine Definition des Begriffs Kurve: die Mengen, die herkömmlicherweise diesen Namen führen, sind von so heterogener Beschaffenheit, daß sie unter keinen vernünftigen Sammelbegriff fallen. Es wäre rationell, ebene Punktmengen jedenfalls nur dann Kurven zu nennen, wenn sie keine inneren Punkte haben; das stetige Bild einer Strecke ist dann im allgemeinen keine Kurve.

in vier Quadrate B_{ikl} derart, daß bei der lexikographischen Anordnung der 64 B_{ikl} nebeneinanderstehende Quadrate benachbart sind usw.

Die Figuren zeigen in der oberen Reihe die Numerierung bei den ersten beiden Teilungen durch hineingesetzte Ziffern, in der unteren Reihe bei den ersten drei Teilungen durch einen die Quadratmittelpunkte verbindenden Streckenzug an.

2	3
1	4

Fig. 41.

22	23	32	33
21	24	31	34
14	13	42	41
11	12	43	44

Fig. 42.

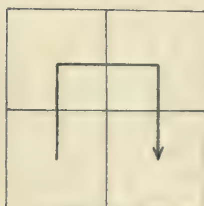


Fig. 43.

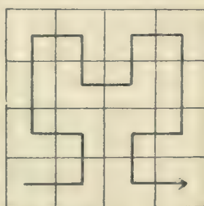


Fig. 44.

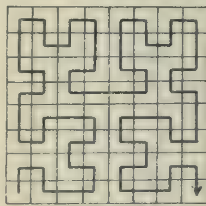


Fig. 45.

Daß bei unbegrenzter Fortsetzung des Verfahrens die vorgeschriebene Numerierung immer möglich ist, bedarf kaum eines Nachweises. Wenn von den vier Teilquadraten eines Quadrats Q das erste Q_1 vorgeschrieben ist, so sind noch zwei Numerierungen möglich, bei denen Q_4 entweder in derselben Horizontal- oder Vertikalreihe wie Q_1 liegt; soll damit Anschluß an ein gegebenes Nachbarquadrat R

Q		R	
2	3		
1	4	1	

Q		R	
1	4	1	
2	3		

Q		R	
2	1		
3	4	1	

Q		R	
3	4	1	
2	1		

Fig. 46.

erreicht werden, so ist von diesen beiden Numerierungen nur eine zulässig und damit auch R_1 bestimmt. (In den Figuren liegt R rechts neben Q; andernfalls denke man sie sich entsprechend gedreht.) Wir haben also für B_1, B_{11}, B_{111}, ... jedesmal die Wahl unter vier und dann noch für B_{44}, B_{444}, ... die Wahl unter zwei möglichen Fällen; die ersten Quadrate können links unten, die letzten rechts unten gewählt werden.

Die Strecke A teilen wir in vier gleiche Teilstrecken $A_1\ A_2\ A_3\ A_4$, die wir einfach von links nach rechts aufeinander folgen lassen, ebenso jedes A_i von links nach rechts in $A_{i1}\ A_{i2}\ A_{i3}\ A_{i4}$ usw.

Ist jetzt x ein Punkt von A, und zwar zuerst ein solcher, der nicht zu den Teilpunkten gehört (x also keine dyadisch rationale Zahl zwischen 0 und 1), so gehört x einer bestimmten Strecke A_i, dann einer bestimmten A_{ik}, dann einer bestimmten A_{ikl} an usw.; x bestimmt also eine gewisse Zahlenfolge $i, k, l, \ldots$ aus den Ziffern 1, 2, 3, 4 derart, daß

$$\{x\} = \mathfrak{D}(A_i, A_{ik}, A_{ikl}, \ldots);$$

denn diese Strecken haben, weil ihre Längen nach 0 konvergieren, nur den einen Punkt x gemein. Die entsprechenden Quadrate $B_i, B_{ik}, B_{ikl}, \ldots$ bilden eine absteigende Folge beschränkter abgeschlossener Mengen, haben also nach dem Cantorschen Satze (S. 230, vgl. auch S. 318) einen und, weil ihre Seiten nach 0 konvergieren, nur einen Punkt y gemein,

$$\{y\} = \mathfrak{D}(B_i, B_{ik}, B_{ikl}, \ldots).$$

Diesen Punkt definieren wir als Bild von x, $y = f(x)$.

Ist x ein Teilpunkt (also dyadisch rational, excl. 0 und 1, für welche die vorige Betrachtung gilt), so gehört er von einer gewissen Teilung an zu zwei benachbarten Strecken, z. B. der Punkt $\frac{3}{16}$ zu

$$A_1,\quad A_{13}\ A_{14},\quad A_{134}\ A_{141},\quad A_{1344}\ A_{1411},\quad A_{13444}\ A_{14111}, \ldots.$$

Er liefert dann zwei solche Zahlenfolgen $i, k, l, \ldots$, von denen die eine mit lauter Vieren, die andere mit lauter Einsen endigt. Trotzdem gibt die obige Bestimmung auch hier nur einen Punkt y, denn die zugehörigen Paare von Quadraten sind benachbart, bilden also Rechtecke, die wieder in absteigender Folge auftreten und wegen ihrer nach 0 konvergierenden Seiten einen einzigen Durchschnittspunkt haben. Die Funktion $y = f(x)$ ist also eindeutig.

Dabei durchläuft y das ganze Quadrat, denn umgekehrt bestimmt jeder Punkt y mindestens ein ihn enthaltendes $B_i, B_{ik}, B_{ikl}, \ldots$ und danach ein $x = \varphi(y)$. Indessen ist diese inverse Funktion mehrdeutig, da benachbarten Quadraten nicht immer benachbarte Strecken entsprechen. Eine nähere, ganz einfache Betrachtung zeigt, daß $\varphi(y)$ einen, zwei oder vier Werte haben kann.

Die Funktion $y = f(x)$ ist stetig: wir zeigen, daß mit $\overline{xx'}$ auch $\overline{yy'}$ nach Null konvergiert. Nehmen wir bereits $\overline{xx'} < \frac{1}{4}$ an und wählen die natürliche Zahl n so, daß $\frac{1}{4^n} > \overline{xx'} \geqq \frac{1}{4^{n+1}}$. Bei der n^{ten} Teilung haben die Teilstrecken die Länge $\frac{1}{4^n}$, die Strecke (x, x') kann also höchstens mit zwei solchen (benachbarten) Teilstrecken

Punkte gemein haben, und ihr Bild höchstens mit zwei benachbarten Teilquadraten von der Seitenlänge $\frac{1}{2^n}$. Folglich ist $\overline{yy'}$ höchstens gleich der Diagonale in einem aus zwei solchen Quadraten gebildeten Rechteck, $\overline{yy'} \leqq \frac{\sqrt{5}}{2^n}$, also

$$\overline{yy'}^2 \leqq \frac{5}{4^n} = \frac{20}{4^{n+1}} \leqq 20 \cdot \overline{xx'}.$$

Damit ist gezeigt, daß das Quadrat stetiges Bild der Strecke sein kann; deuten wir x als Zeit, so kann ein Punkt bei stetiger Bewegung eine Bahn beschreiben, die eine ganze Fläche ausfüllt.[1]

Man kann diese Betrachtung dadurch modifizieren, daß man die Quadratseiten statt in 2, in 3 oder mehr Teile teilt (Teilung in eine ungerade Anzahl führt sogar zu einer relativ einfachen arithmetischen Darstellung der Funktion $f(x)$, auf die wir nicht eingehen wollen); auch ist evident, daß man an Stelle des Quadrats den drei- oder mehrdimensionalen Würfel setzen kann.

Ganz anders liegen die Verhältnisse, wenn wir B als umkehrbar eindeutiges, stetiges Bild der Strecke A voraussetzen.[2] Nach § 1, VIII ist dann auch A stetiges Bild von B; jeder zusammenhängenden Teilmenge der einen Menge entspricht eine zusammenhängende Teilmenge der andern. Da A durch Weglassung eines Punktes (mit Ausnahme eines der beiden Endpunkte) unzusammenhängend wird, so gilt dasselbe von B; diese Menge hat also, als Menge in einem mindestens zweidimensionalen Raume betrachtet, keine inneren Punkte (S. 333). Sie wird als ungeschlossene Jordansche Kurve oder auch als Jordanscher Kurvenbogen bezeichnet; zwei solche Kurvenbögen sind eindeutige stetige Bilder von einander.

Das eineindeutige stetige Bild B eines Polygons oder Kreises[3] A heißt eine geschlossene Jordansche Kurve; auch hier ist A stetiges Bild von B, und B ist beschränkt. B kann keine lineare Menge sein, da sie sonst schon durch Tilgung eines geeigneten (zwischen zwei andern liegenden) Punktes unzusammenhängend würde;

---

[1] Projiziert man das Quadrat auf eine seiner Seiten, so erhält man eine Bewegung, bei der eine Strecke beschrieben, aber jeder ihrer Punkte $\aleph$-mal durchlaufen wird ($\aleph$ die Mächtigkeit des Kontinuums).

[2] Bei den Funktionen (1) ist dann zu verlangen, daß die sämtlichen Gleichungen $f_1(x') = f_1(x), \ldots, f_n(x') = f_n(x)$ nur für $x' = x$ zusammen bestehen.

[3] Die Gleichungen der vorigen Anmerkung dürfen dann nur für $x' = x$ und außerdem für die Intervallendpunkte $x = 0$, $x' = 1$ zusammen bestehen. Man kann dann in ersichtlicher Weise die Funktionen $f_1(x), \ldots, f_n(x)$ als periodische Funktionen (mit der Periode 1) für alle reellen x fortsetzen; sie müssen überall stetig und für $0 \leqq x < 1$ eindeutig umkehrbar sein.

sie gehört also einem E_n ($n \geqq 2$) an. Da B durch Tilgung zweier Punkte unzusammenhängend wird, hat sie keine inneren Punkte. Jeder abgeschlossenen und jeder zusammenhängenden Teilmenge von A entspricht eine ebensolche von B und umgekehrt, und damit sind die Voraussetzungen verwirklicht, auf die wir im Falle der Ebene den Beweis des Jordanschen Satzes (S. 354) gegründet haben. Für eine in der Ebene E liegende geschlossene Jordansche Kurve B zerfällt also $E - B$ in zwei zusammenhängende Gebiete, deren gemeinsame Grenze B ist.

Für einen Jordanschen Kurvenbogen B ist $E - B$ ein zusammenhängendes Gebiet mit der Grenze B.

Wir heben ausdrücklich hervor, daß nicht jede geschlossene Kurve eine Jordansche ist. Der in der untenstehenden Figur veranschaulichte Streckenzug

$$S = (a_0, a_1, a_2, a_3, a_4, b_1, c_1, c_2, b_2, b_3, c_3, c_4, b_4, \ldots),$$

bestehend aus abzählbar unendlich vielen Strecken, deren Endpunkte die rechtwinkligen Koordinaten haben:

Punkt:	a_0	a_1	a_2	a_3	a_4	b_n	c_n
Abszisse:	0	0	0	2	2	$\frac{1}{n}$	$\frac{1}{n}$
Ordinate:	0	1	2	2	0	0	1

ist eine geschlossene Kurve; das Innengebiet ist die obere Hälfte vom Innern des Quadrats $a_0\, a_2\, a_3\, a_4$ nebst den nach unten angefügten rechteckigen Zungen (wie $c_2\, b_2\, b_3\, c_3$), das Außengebiet das Äußere des Quadrats mit den nach oben reichenden Zungen (wie $b_1\, c_1\, c_2\, b_2$), und jeder Punkt von S ist Grenzpunkt von beiden Gebieten. Aber S ist keine Jordansche Kurve, da sie nach Tilgung der beiden Punkte a_0, a_1 oder, allgemeiner, beliebig vieler Punkte der Strecke $T = (a_0, a_1)$ zusammenhängend bleibt. Denn $S - T$, der Streckenzug $(a_1, a_2, \ldots)$ ohne den Punkt a_1, ist noch zusammenhängend und hat die Punkte von T zu Häufungspunkten, also ist $(S - T)_\alpha = S$ und jede Menge zwischen $S - T$ und S zusammenhängend.[1]

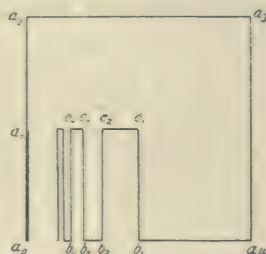


Fig. 47.

[1] Dies steht damit in Beziehung, daß die Punkte der Strecke T, abgesehen von a_1, von dem Innengebiet aus nicht erreichbar sind (S. 347). A. Schoenflies hat, allerdings mit einem etwas spezieller erklärten Erreichbarkeitsbegriff, bewiesen, daß eine geschlossene Kurve dann und nur dann eine Jordansche ist, wenn alle ihre Punkte sowohl vom inneren wie vom äußeren Gebiet aus erreichbar sind.

Ändert man das Beispiel so ab, daß man c_n die Koordinaten $\frac{1}{n}, \frac{1}{n}$ gibt, so wird S eine Jordansche Kurve, die man (indem man (b_1, c_1, c_2, b_2) durch (b_1, b_2) usw. ersetzt) umkehrbar eindeutig und stetig auf das Quadrat $(a_0, a_2, a_3, a_4, a_0)$ abbilden kann.

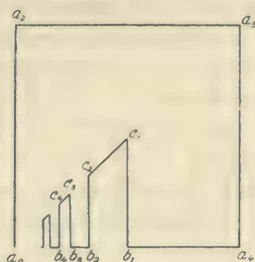


Fig. 48.

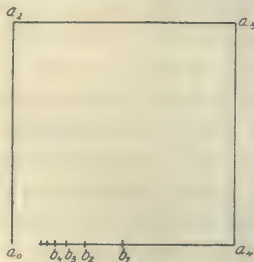


Fig. 49.

Es gibt Jordansche Kurven und Kurvenbögen, die positives Flächenmaß haben.[1] Wir zeigen dies, indem wir durch die S. 326 (Fig. 12) konstruierte punkthafte Menge, der man ja positives Flächenmaß geben kann, einen Kurvenbogen legen. In das Quadrat B setzen wir vier, ein Kreuz freilassende Quadrate, die wir jetzt (in der Reihenfolge links unten, links oben, rechts unten, rechts oben) mit B_1, B_3, B_5, B_7 bezeichnen; verbinden wir noch den rechten oberen Eckpunkt eines Quadrats mit dem linken unteren

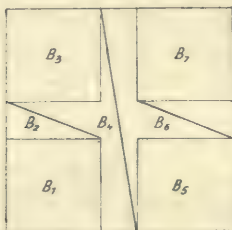


Fig. 50.



Fig. 51.

des nächsten, so erhalten wir drei Strecken B_2, B_4, B_6, die bis auf die Endpunkte in dem kreuzförmigen Gebiet verlaufen.[2] Um dies Verfahren fortzusetzen, verabreden wir, daß $p, q, r, \ldots$ die Ziffern

---

[1] Daß eine Jordansche Kurve ein Innengebiet mit positivem Flächenmaß begrenzt, versteht sich von selbst; aber hier ist gemeint, daß die Kurve selbst, obwohl ohne innere Punkte, positives Flächenmaß haben kann.

[2] Unter den Strecken sind abgeschlossene Strecken, unter den Quadraten abgeschlossene Quadratflächen verstanden.

1, 3, 5, 7 durchlaufen sollen, t die Ziffern 2, 4, 6, endlich $i, k, l, \ldots$ alle Ziffern 1, 2, ..., 7. In jedes Quadrat B_p werden vier kleinere Quadrate B_{pq}, mit freibleibendem Mittelkreuz, nebst den drei Verbindungsstrecken B_{pt} gesetzt, in jedes Quadrat B_{pq} vier Quadrate B_{pqr} mit den drei Verbindungsstrecken B_{pqt} usw. Außerdem teilen wir jede Strecke B_t in sieben gleiche Teilstrecken B_{tk}, die wir, vom Schnittpunkt mit B_{t-1} nach dem Schnittpunkte mit B_{t+1} zählend, mit $B_{t1}, B_{t2}, \ldots, B_{t7}$ bezeichnen, ebenso jede Strecke B_{pt} und B_{tk} in sieben Strecken B_{ptl} und B_{tkl} usw. Der Durchschnitt

$$Y = \mathfrak{D}(\mathfrak{S} B_i, \mathfrak{S} B_{ik}, \mathfrak{S} B_{ikl}, \ldots),$$

der die punkthafte perfekte Menge

$$\mathfrak{D}(\Sigma B_p, \Sigma B_{pq}, \Sigma B_{pqr}, \ldots)$$

umfaßt, ist dann ein Jordanscher Kurvenbogen. Um ihn auf die Einheitsstrecke A abzubilden, teilen wir diese in sieben gleiche Teilstrecken A_i (von links nach rechts: $A_1, \ldots, A_7$), jede davon wieder in sieben Teilstrecken A_{ik} usw. Nun beachten wir: die Mengen B_i sind paarweise fremd, bis auf zwei lexikographisch benachbarte (B_i und B_{i+1}), die genau einen Punkt gemein haben, ebenso die Mengen B_{ik} bis auf zwei lexikographisch benachbarte (B_{ik} und $B_{i(k+1)}$ oder B_{i7} und $B_{(i+1)1}$) usw. Bezeichnet man die Diagonalenlänge des Quadrats B mit $\frac{\delta}{2}$, so sind die Breiten der Mengen $B_i, B_{ik}, B_{ikl}, \ldots$ kleiner als $\frac{\delta}{2}, \frac{\delta}{2^2}, \frac{\delta}{2^3}, \ldots$. Lassen wir demnach dem Durchschnittspunkt x einer Folge $A_i, A_{ik}, A_{ikl}, \ldots$ den Durchschnittspunkt y der Folge $B_i, B_{ik}, B_{ikl}, \ldots$ als Bild entsprechen, so ist diese Beziehung umkehrbar eindeutig, auch für die Punkte, die zwei solchen, lexikographisch benachbarten Mengenfolgen angehören (z. B. gehört der linke untere Eckpunkt des Quadrats B_3 den Mengenfolgen $B_2, B_{27}, B_{277}, \ldots$ und $B_3, B_{31}, B_{311}, \ldots$ an, entspricht also dem linken Endpunkt der Strecke A_3). Schließlich ist die Funktion $y = f(x)$ stetig; denn ist $\frac{1}{7^n} > \overline{xx'} \geqq \frac{1}{7^{n+1}}$, so gehören y, y' einer oder zwei benachbarten Mengen $B_{ikl} \ldots$ mit n Ziffern an, ihre Entfernung ist höchstens gleich der Breitensumme zweier solcher Mengen, also $\overline{yy'} < \frac{2\delta}{2^n}$ und

$$\overline{yy'}^3 < \frac{8\delta^3}{8^n} < 56\,\delta^3\,\overline{xx'}.$$

Um die Bedeutung des Jordanschen Satzes richtig zu beurteilen, bemerken wir: wenn zwei ganze Ebenen E_x, E_y eindeutige stetige Bilder von einander sind, so ist für diese Abbildung der Jordansche Satz und noch vieles darüber hinaus eine

Selbstverständlichkeit. Jeder abgeschlossenen Menge der einen Ebene entspricht eine abgeschlossene der andern,[1] jedem Gebiet der einen ein Gebiet der andern; Zusammenhang und Komponentenzahl bleiben ungeändert. Das Bild eines k-fach zusammenhängenden Gebiets ist ein ebensolches, das Bild einer geschlossenen Kurve wieder eine geschlossene Kurve u. dgl. Dies alles gilt entsprechend auch für umkehrbar eindeutige, beiderseits stetige Abbildung beliebiger euklidischer Räume.

Viel schwieriger ist es, Sätze von folgendem Charakter zu beweisen: das umkehrbar eindeutige, stetige Bild einer Menge A von gewisser Beschaffenheit ist eine Menge B von gewisser Beschaffenheit; wobei also nur eine in A selbst (nicht in der ganzen Ebene E_x) stetige, eineindeutige Funktion $f(x)$ vorausgesetzt wird und unter Umständen das Problem hinzutritt, auch die Stetigkeit der inversen Funktion $\varphi(y)$ zu beweisen (vgl. den Satz § 1, VIII und in diesem Paragraphen den Satz II). Von diesem Charakter ist der Jordansche Satz, der nur eine auf dem Kreise stetige, umkehrbar eindeutige Funktion als gegeben annimmt, und von gleicher Art werden auch die obigen Sätze über Invarianz des Gebietes, des k-fach zusammenhängenden Gebiets,[2] der geschlossenen Kurve usw., wenn man eben nur diese Mengen selbst, nicht die ganze Ebene, einer eineindeutigen stetigen Abbildung unterwirft. Man pflegt die Betrachtung solcher Eigenschaften und Beziehungen von Punktmengen, die bei umkehrbar eindeutiger und (beiderseits) stetiger Abbildung invariant bleiben, als Analysis situs zu bezeichnen; schreibt man Mengen, die eindeutige stetige Bilder von einander sind, denselben geometrischen Typus zu, genau wie man ähnlich geordneten Mengen denselben Ordnungstypus zuschreibt, so ist Analysis situs die Theorie der geometrischen Typen, deren Aufzählung, Klassifikation usw. ihr obliegt. Aber diese Theorie befindet sich nach unserer Meinung noch durchaus nicht in dem erwünschten Zustande von Einfachheit und Vollständigkeit; wir können hier nicht sehr tief in sie eindringen und werden auch von den oben behaupteten Invarianzen nur einen Teil beweisen, wobei sich Gelegenheit zu Anwendungen des Jordanschen Satzes ergeben wird.

Die wichtigste dieser Anwendungen ist der Beweis der Invarianz der Dimensionenzahl gegenüber umkehrbar eindeutiger, beider-

---

[1] Daß einer abgeschlossenen Menge $\leqq E_x$ eine abgeschlossene $\leqq E_y$ entspricht, folgt aus der Stetigkeit der Funktion $\varphi(y)$ nach § 1, III; aus der Stetigkeit von $f(x)$ würde es nach § 1, V nur für beschränkte abgeschlossene Mengen folgen.

[2] Hierbei sind für unbeschränkte Gebiete besondere Vereinbarungen zu treffen.

seits stetiger Abbildung. Wir erinnern zunächst daran, daß gerade Linie, Ebene, Raum, Strecke, Quadrat, Würfel usw. von gleicher Mächtigkeit $\aleph$ (der des Kontinuums) sind, also umkehrbar eindeutig auf einander abgebildet werden können. In Kap. III, § 5 haben wir eine solche (im Prinzip von G. Cantor herrührende) Abbildung des Quadrats $(0 < x, y \leqq 1)$ auf die Strecke $(0 < z \leqq 1)$ mit Hilfe der Formeln

$$z = [c_1, c_2, c_3, c_4, \ldots] = (\tfrac{1}{2})^{c_1} + (\tfrac{1}{2})^{c_1+c_2} + (\tfrac{1}{2})^{c_1+c_2+c_3} + \cdots,$$
$$x = [c_1, c_3, c_5, \ldots], \quad y = [c_2, c_4, c_6, \ldots]$$

angegeben. Aber diese Funktionen $x = f(z)$, $y = g(z)$, $z = \varphi(x, y)$ sind keineswegs stetig. Z. B. für $z_0 = \frac{1}{2} = [2, 1, 1, 1, \ldots]$ wird $x_0 = \frac{1}{2}$, $y_0 = 1$. Für $z > \frac{1}{2}$, $z = [1, c_2, c_3, \ldots]$ wird $x = [1, c_3, c_5, \ldots]$, $y = [c_2, c_4, c_6, \ldots]$. Läßt man hier z abnehmend nach $\frac{1}{2}$ konvergieren, so wächst c_2 über alle Grenzen, während $c_3, c_4, \ldots$ ganz beliebig variieren können; dabei konvergiert y nach 0, während x beliebig zwischen $\frac{1}{2}$ (exkl.) und 1 (inkl.) oszillieren kann, beide Funktionen sind an der Stelle $z = \frac{1}{2}$ unstetig. Ähnliches wiederholt sich in jedem noch so kleinen Intervall für z und gilt auch für die inverse Funktion $z = \varphi(x, y)$.

Andererseits zeigt die Peanosche Kurve, daß das Quadrat sogar eindeutiges, stetiges Bild der Strecke sein kann; hier war aber die inverse Funktion nicht eindeutig. Die Cantorsche und die Peanosche Abbildung zusammen legen die Vermutung nahe, daß der ins Schwanken geratene Dimensionsbegriff sich wieder herstellen werde, wenn man umkehrbar eindeutige, beiderseits stetige Abbildung verlangt. Wir formulieren diese Vermutung präziser in dem folgenden Satz von der Invarianz der Dimensionenzahl:

I. Die Menge A gehöre einem m-dimensionalen euklidischen Raume an, die Menge B einem n-dimensionalen $(n > m)$ und habe innere Punkte. Wenn dann A und B umkehrbar eindeutig auf einander abgebildet sind, so ist jedenfalls A kein stetiges Bild von B.

Unsere Hilfsmittel setzen uns nur in den Stand, den Satz für die niedrigsten Fälle $m = 1$ und $m = 2$ zu beweisen.

Für $m = 1$ ist A eine lineare Menge oder eine Menge reeller Zahlen. Seien b, b_1, b_2 drei innere Punkte von B und a, a_1, a_2 ihre Bilder; der Größe nach sei $a_1 < a < a_2$. Die Menge $B - \{b\}$ ist noch zusammenhängend, die Menge $A - \{a\}$ aber nicht, also kann A nicht stetiges Bild von B sein.

Für $m = 2$ ist A eine ebene Menge. Betrachten wir einen inneren Punkt von B und eine in B enthaltene Umgebung desselben; indem wir uns auf die Betrachtung dieser beschränken, können wir B selber als Inneres einer drei- oder mehrdimensionalen Kugel

annehmen. Indem wir B etwa mit einer Schar paralleler Ebenen schneiden, können wir unabzählbar viele in B liegende, einander nicht treffende Kreislinien $Q, Q_1, Q_2, \ldots$ angeben. Ihre Bilder $P, P_1, P_2, \ldots$ würden, wenn A stetiges Bild von B wäre, unabzählbar viele paarweise fremde Jordansche Kurven einer Ebene sein. Es ist unmöglich, daß keine von ihnen die übrigen trenne (S. 346); möge also etwa P_1 im Innern und P_2 im Äußern von P liegen. Die Menge $B-Q$ ist offenbar noch zusammenhängend, die Menge $A-P$ aber nicht, denn sie zerfällt in zwei nichtverschwindende Relativgebiete A_1 (Menge ihrer Punkte im Innern von P) und A_2 (Menge ihrer Punkte im Äußern von P). Also kann A nicht stetiges Bild von B sein.

Die Übertragung dieses Beweises auf $m > 2$ würde von der Übertragung des Jordanschen Satzes auf den m-dimensionalen Raum abhängen. Wir verweisen den Leser in dieser Hinsicht auf die Arbeiten von L. E. J. Brouwer.

Daß, unter den Voraussetzungen von I, B sehr wohl stetiges Bild von A sein kann, zeigt eine leichte Modifikation der Peanoschen Abbildung. Behält man dort von den Urbildern $a = \varphi(b)$ eines Quadratpunktes b nur ein einziges bei, so wird das Quadrat B stetiges und umkehrbar eindeutiges Bild einer gewissen linearen Punktmenge A'.

Ist, unter den Voraussetzungen von I, A abgeschlossen und beschränkt, so kann auch B kein stetiges Bild von A sein, da nach § 1, VIII sonst auch A stetiges Bild von B wäre. Bei eineindeutiger Abbildung der Strecke auf das Quadrat (beide abgeschlossen gedacht) ist also keine Menge stetiges Bild der andern.

Um in der Theorie der Abbildung ebener Punktmengen noch einen Schritt weiterzugehen, betrachten wir eine abgeschlossene Kreisfläche mit dem Mittelpunkt a und dem Radius τ, die umkehrbar eindeutig und stetig auf eine ebene Punktmenge abgebildet wird. P_ϱ sei die Kreislinie mit dem Mittelpunkt a und dem Radius $\varrho \leqq \tau$; ihr Bild ist eine Jordansche Kurve $Q_\varrho = F(P_\varrho)$, die wegen der Stetigkeit für hinlänglich kleines ϱ in beliebiger Nähe des Bildpunktes $b = f(a)$ verläuft: wir schließen daraus, daß Q_τ für hinlänglich kleines $\varrho < \delta$ außerhalb (im Außengebiet) von Q_ϱ liegt.[1] Andererseits dürfen, für $\varrho < \sigma < \tau$, Q_σ und Q_τ durch Q_ϱ nicht getrennt werden; denn P_σ und P_τ lassen sich durch eine zusammenhängende Menge (etwa das Stück eines Radius) verbinden, die P_ϱ nicht trifft, und dasselbe gilt für ihre Bilder. Q_σ und Q_τ liegen also auf derselben Seite von Q_ϱ, beide außerhalb oder beide innerhalb;

[1] Man ziehe um b als Mittelpunkt einen kleinen Kreis, außerhalb dessen Q_τ liegt; Q_ϱ liegt aber schließlich innerhalb dieses Kreises.

für $\varrho < \sigma < \delta$ liegt also auch Q_σ außerhalb Q_ϱ. Aus demselben Grunde liegen b und Q_ϱ auf derselben Seite von Q_σ. Läge nun b außerhalb aller jener Jordanschen Kurven, so würde auch Q_ϱ außerhalb Q_σ liegen, und wir hätten unabzählbar viele Jordansche Kurven, deren jede außerhalb der übrigen liegt; da dies nicht sein darf (S. 346), so muß es mindestens ein Q_ϱ geben, innerhalb dessen b liegt. Wenn wir bei dieser Überlegung τ durch einen kleineren Wert ersetzen, so sieht man, daß es solche Q_ϱ mit beliebig kleinem ϱ gibt.

Nach dieser Vorbetrachtung zeigen wir, daß die inneren Punkte einer solchen Kreisfläche auf innere Punkte der Bildmenge abgebildet werden. Mit etwas geänderter Bezeichnung sei $A = P + M$ eine Kreisfläche, P die sie begrenzende Kreislinie, M das Innere; durch eine in A stetige, umkehrbar eindeutige Funktion $f(x)$ erhalte man daraus die ebenen Punktmengen

$$B = F(A), \quad Q = F(P), \quad N = F(M); \quad B = Q + N.$$

B ist beschränkt und abgeschlossen, ihr Rand B_r die Grenze des Gebietes $E - B = G + G' + \ldots$, dessen Komponenten wir hiermit ersichtlich machen; werden deren Grenzen mit $H, H', \ldots$ bezeichnet und $S = \mathfrak{S}(H, H', \ldots)$ gesetzt, so ist (S. 332) $B_r = S_\alpha$. Wir zeigen zunächst, daß N und H keinen Punkt gemein haben. Wäre $b = f(a)$ ein solcher Punkt, also a ein Punkt von M, so betrachte man eine in M enthaltene abgeschlossene Kreisfläche mit dem Mittelpunkt a und die in ihr verlaufenden konzentrischen Kreise P_ϱ; unter deren Bildern gibt es, nach der zuvor angestellten Überlegung, eine beliebig nahe an b verlaufende, der Menge N angehörige Jordansche Kurve Q_ϱ, die b im Innern enthält. Da b Häufungspunkt des zusammenhängenden Gebietes G ist und man ϱ so klein wählen kann, daß G nicht mehr innerhalb Q_ϱ liegt, während doch gewiß Punkte von G innerhalb Q_ϱ liegen, so müßte G die Kurve Q_ϱ treffen, was wegen $Q_\varrho \leqq N \leqq B$ ein Widerspruch ist. Also haben H und N keinen Punkt gemein, und da H dem Rande von $B = Q + N$ angehört, ist $H \leqq Q$. Dasselbe gilt für $H', H'', \ldots$, also folgt $S \leqq Q$, $S_\alpha \leqq Q$ (Q ist abgeschlossen), $B_r \leqq Q$ und $B_i \geqq N$, die Punkte von N oder die Bildpunkte des Kreisinnern sind innere Punkte des Bildes B.

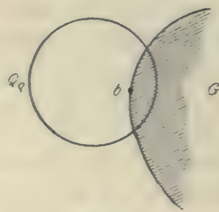


Fig. 52.

Hieraus folgt unmittelbar:

II. (Satz von E. Jürgens.) Ist die ebene Punktmenge B umkehrbar eindeutiges, stetiges Bild des ebenen Gebietes A, so ist auch B ein Gebiet und A stetiges Bild von B.

Denn umgibt man einen beliebigen Punkt a von A mit einer abgeschlossenen Kreisfläche $A' \leqq A$, so wird sein Bildpunkt b innerer Punkt von B', also auch von B. Da nun auch jedem Relativgebiet von A (nämlich jedem Gebiet $\leqq A$) wieder ein Relativgebiet von B entspricht, so ist nach § 1, III auch A stetiges Bild von B.

Im Anschluß daran läßt sich leicht die Invarianz gewisser anderer Figuren feststellen, z. B.: eine Jordansche Kurve und ihr Inneres geht bei umkehrbar eindeutiger, stetiger Abbildung (auf eine ebene Punktmenge) wieder in eine Jordansche Kurve und ihr Inneres über. Denn ist $P+M$ eine Jordansche Kurve und ihr Inneres, $Q+N$ eineindeutiges, stetiges Bild dieser Menge, so ist Q eine Jordansche Kurve, N ein zusammenhängendes Gebiet; ferner seien N_1, N_2 das Innen- und Außengebiet von Q, in irgendwelcher Reihenfolge. N muß ganz dem einen oder andern dieser Gebiete angehören, etwa $N \leqq N_1$; da nun $Q+N$ abgeschlossen ist (§ 1, V), so ist $\mathfrak{D}(N_1, Q+N) = N$ in N_1 abgeschlossen, $N_1 - N$ Relativgebiet von N_1 oder absolutes Gebiet, und die Zerlegung $N_1 = N + (N_1 - N)$ in zwei Gebiete zeigt wegen des Zusammenhangs von N_1, daß der zweite Summand verschwinden muß. Also ist $N = N_1$, N ist eins der von Q übriggelassenen Gebiete, und zwar das Innengebiet, da $Q+N$ beschränkt ist. — Für eine Jordansche Kurve und ihr Äußeres gilt der Satz nicht, weil die Bildmenge $Q+N$ dann nicht abgeschlossen zu sein braucht; z. B. läßt sich ja durch Inversion ein Kreis und sein Äußeres eineindeutig stetig auf denselben Kreis und sein Inneres ohne den Mittelpunkt abbilden.

Ganz ähnlich zeigt man, daß zwei Jordansche Kurven, von denen eine im Innern der andern liegt, und das von ihnen begrenzte ringförmige Zwischengebiet bei eineindeutiger stetiger Abbildung wieder in eine solche Figur übergehen. Denn wenn $P_1 + P_2 + M$ in $Q_1 + Q_2 + N$ übergeht, so ist wieder N ein zusammenhängendes Gebiet, während die beiden Jordanschen Kurven Q_1, Q_2 drei zusammenhängende Gebiete N_1, N_2, N_3 übrig lassen (S. 338, 346); wie zuvor ist N mit einem von diesen identisch. Und zwar muß N das Zwischengebiet sein, da beide Jordansche Kurven Q_1, Q_2 seine Grenze bilden, und weil N beschränkt ist, muß eine dieser Kurven im Innern der andern liegen.

III. Das ebene, beschränkte, umkehrbar eindeutige, stetige Bild eines ebenen, beschränkten, k-fach zusammenhängenden Gebietes ist wieder ein k-fach zusammenhängendes Gebiet.

Die fraglichen Gebiete[1] seien A, B; wir haben zu zeigen, daß

[1] Wir können sie derselben Ebene E angehörig annehmen.

ihre (beschränkten) Grenzen C, D gleichviele Komponenten haben. Zunächst ergibt sich nun zwischen den Punkten c, d dieser Grenzen eine, im allgemeinen beiderseits mehrdeutige Beziehung folgendermaßen. Bestimmen wir zu einem Grenzpunkt c eine nach ihm konvergente, dem Gebiet A angehörige Menge $X = \{a_1, a_2, \ldots\}$. Die Bildmenge $Y = \{b_1, b_2, \ldots\}$, dem beschränkten Gebiete B angehörig, hat mindestens einen Häufungspunkt, der zu B_α gehört; er kann aber kein Punkt b des Gebietes B selbst sein, da diesem ein Punkt a von A entsprechen würde, der (weil auch A nach II stetiges Bild von B ist) Häufungspunkt von X sein müßte. Also ist der genannte Häufungspunkt ein Punkt d der Grenze D, und für eine geeignete Teilfolge ist $c = \lim a_p$, $d = \lim b_p$. Mit andern Worten: jedem Punkt c entspricht mindestens ein Punkt $d = f(c)$ derart, daß gleichzeitig eine Menge $X \leqq A$ nach c und ihre Bildmenge $Y \leqq B$ nach d konvergiert, und umgekehrt entspricht einem Punkt d in gleicher Weise mindestens ein Punkt $c = \varphi(d)$; aber beide Funktionen sind im allgemeinen mehrdeutig.[1] Nun ist aber das Charakteristische, daß die Komponenten beider Grenzen einander eineindeutig zugeordnet werden, auf Grund folgender zunächst zu beweisender Tatsache: wenn c_1, c_2 zwei verschiedenen Komponenten von C angehören, so gehören irgend zwei entsprechende Punkte d_1, d_2 verschiedenen Komponenten von D an. In der Tat gibt es nach dem Satze VII in Kap. VIII, § 11 ein in A verlaufendes Polygon P, das c_1 und c_2 trennt; es seien M_1, M_2 die Komponenten von $E - P = M_1 + M_2$, und c_1 liege in M_1, c_2 in M_2. Bei der Zerlegung

$$A - P = A_1 + A_2 = \mathfrak{D}(A, M_1) + \mathfrak{D}(A, M_2)$$

sind die beiden Gebiete A_1, A_2 zusammenhängend (S. 338), und dies würde nach S. 346 auch richtig bleiben, wenn man das Polygon durch eine geschlossene Kurve, z. B. eine Jordansche ersetzt. Das Bild von P ist eine Jordansche Kurve Q in B, und diese bewirkt also die analoge Zerlegung

$$B - Q = B_1 + B_2 = \mathfrak{D}(B, N_1) + \mathfrak{D}(B, N_2),$$

wo N_1, N_2 die Komponenten von $E - Q = N_1 + N_2$ sind; überdies folgt, daß die Zerlegungen $B_1 + B_2 = F(A_1) + F(A_2)$ in zwei zusammenhängende Gebiete identisch sein müssen, und bei passender Nume-

---

[1] Bildet man das Innere eines Kreises vom Radius 1 auf das eines andern durch die Gleichungen in Polarkoordinaten

$$r' = r, \; \varphi' = \varphi + \frac{r}{1 - r}$$

ab, so entsprechen jedem Punkte der einen Kreisperipherie alle Punkte der andern.

rierung sind also B_1, B_2 die Bilder von A_1, A_2. Endlich seien nun X_1, X_2, Y_1, Y_2 Mengen, die nach c_1, c_2, d_1, d_2 konvergieren. X_1 liegt, bis auf endlich viele Punkte, auf derselben Seite von P wie c_1, also in M_1 und in A_1; bei eventueller Weglassung von endlich vielen Punkten können wir also $X_1 \leqq A_1$, $X_2 \leqq A_2$ annehmen. Dann liegen die Bildmengen Y_1, Y_2 in B_1, B_2, also in N_1, N_2. Der Limes d_1 von Y_1 liegt in $N_{1\alpha} = N_1 + Q$, also in N_1, da Q und D einander nicht treffen; ebenso liegt d_2 in N_2, und hiermit ist

$$D = \mathfrak{D}(D, N_1) + \mathfrak{D}(D, N_2) = D_1 + D_2$$

in zwei Relativgebiete oder zwei abgeschlossene Mengen zerlegt, deren eine d_1, die andere d_2 enthält. Demnach gehören d_1 und d_2 verschiedenen Komponenten von D an.

Wenn also c_1, c_2 verschiedenen Komponenten von C angehören, so d_1, d_2 verschiedenen Komponenten von D, und umgekehrt (da auch A stetiges Bild von B ist). Damit sind also die Komponenten beider Grenzen in eineindeutige Beziehung gebracht und der Satz III bewiesen, in dem k endlich oder $\aleph_0$ oder $\aleph$ sein kann. Auf die Modifikationen (Adjunktion eines Punktes ∞), die notwendig sind, um den Satz auch für unbeschränkte Gebiete aufrechtzuerhalten, gehen wir nicht ein.

§ 3. Unstetige Funktionen.

Je nachdem die in A definierte Funktion $f(x)$ im Punkte a stetig ist oder nicht, wird a ein Stetigkeitspunkt oder Unstetigkeitspunkt genannt; die Menge der Stetigkeitspunkte heiße C, die der Unstetigkeitspunkte D, also $A = C + D$. Bisher behandelten wir den Fall, daß $f(x)$ in A überall stetig ist ($C = A$, $D = 0$); das andere Extrem wäre, daß $f(x)$ überall unstetig ist ($C = 0$, $D = A$), z. B. die Dirichletsche Funktion einer reellen Variablen, die an den rationalen Stellen $= 1$ und an den irrationalen $= 0$ ist (S. 267). Dazwischen liegen die Fälle, wo sowohl Stetigkeits- als Unstetigkeitspunkte auftreten. Wenn die Menge der Stetigkeitspunkte in A dicht ist, so heißt die Funktion in A (höchstens) punktweise unstetig; dazu ist auch der Fall einer in A stetigen Funktion zu rechnen.

Wenn $f(x)$ in a stetig ist, so gibt es für jede Umgebung V_b des Bildpunktes $b = f(a)$ eine Umgebung U_a, deren Bild in V_b liegt, oder, wenn wir metrische Räume voraussetzen (wenigstens soll die Bildmenge B metrisch sein), so gibt es zu jedem positiven ε ein U_a, für dessen sämtliche[1] Punkte x die Entfernung $\overline{by} < \varepsilon$ ist. Für

[1] Wir sehen wieder A, B als die zugrunde liegenden Räume an.

irgend zwei zu U_a gehörige Punkte x, x' ist dann die Entfernung der Bilder $\overline{yy'} < 2\varepsilon$, und wir können daher die Stetigkeit in a auch so formulieren, daß zu jedem ε ein U_a existiert, dessen sämtliche Punktpaare Bilder mit einer Entfernung $< \varepsilon$ haben. Bezeichnen wir die Menge derjenigen Punkte a von A, die für ein bestimmtes ε eine solche Umgebung U_a haben, mit $A(\varepsilon)$, so ist die Menge C der Stetigkeitspunkte der Durchschnitt aller $A(\varepsilon)$, wenn ε alle positiven Zahlen durchläuft; da aber offenbar $A(\varepsilon) \leqq A(\varepsilon')$ für $\varepsilon < \varepsilon'$, so genügt es, ε eine abnehmende, nach Null konvergente Folge durchlaufen zu lassen, z. B.

$$C = \mathfrak{D}\left(A(1),\ A(\tfrac{1}{2}),\ A(\tfrac{1}{3}),\ \ldots\right).$$

Nun ist unmittelbar ersichtlich, daß jedes $A(\varepsilon)$ ein Gebiet (Relativgebiet in A) ist; denn ist a ein Punkt von $A(\varepsilon)$ und U_a seine wie oben bestimmte Umgebung, so ist für einen Punkt x von U_a eine Umgebung $U_x \leqq U_a$ angebbar, deren Punktpaare ebenfalls Bilder mit einer Entfernung $< \varepsilon$ haben, d. h. x ist selber Punkt von $A(\varepsilon)$ oder a ein innerer Punkt von $A(\varepsilon)$. Die vorangehende Formel für C, und die für das Komplement $D = A - C$, ergeben also den Satz:

Die Menge der Stetigkeitspunkte einer in A definierten Funktion ist (in bezug auf den Raum A) Durchschnitt einer Folge von Gebieten, die Menge der Unstetigkeitspunkte Summe einer Folge abgeschlossener Mengen.

Also C ist ein G_δ, D ein F_σ in unserer gewohnten Bezeichnung; will man sich, statt auf A, auf einen Raum $E_x \geqq A$ beziehen, so ist C eine Menge der Form $\mathfrak{D}(A, G_\delta)$, D eine Menge der Form $\mathfrak{D}(A, F_\sigma)$.

Für eine punktweise unstetige Funktion schließen sich hieran die Betrachtungen von Kap. VIII, § 9 am Ende. Danach ist in diesem Fall C in A dicht, also D in bezug auf A eine Menge erster Kategorie, d. h. die Summe einer Folge von Mengen, die in A nirgendsdicht sind. Ist daher A in bezug auf sich selbst von zweiter Kategorie, etwa eine Menge von der Form G_δ in einem vollständigen Raume (z. B. ein Gebiet oder eine abgeschlossene Menge im euklidischen Raume), so ist C ebenfalls von zweiter Kategorie in bezug auf A; die Mengen der Stetigkeits- und der Unstetigkeitspunkte sind also von ganz verschiedenem Charakter. Z. B. kann eine Funktion einer reellen Variablen wohl an den rationalen Stellen unstetig, an den irrationalen stetig sein, nicht aber umgekehrt; denn die Menge der rationalen Zahlen ist ein F_σ und von erster Kategorie, die der irrationalen Zahlen ein G_δ und von zweiter Kategorie in bezug auf die Menge der reellen Zahlen. Als Beispiel definiere man etwa, für irrationales x, $f(x) = 0$; für

rationales $x = \frac{p}{q}$, wo p, q teilerfremd und q positiv ist, $f(x) = \frac{1}{q}$. Ein anderes Beispiel ist das von **Riemann** angegebene

$$f(x) = \frac{(x)}{1^2} + \frac{(2x)}{2^2} + \frac{(3x)}{3^2} + \cdots,$$

wo (x) der Überschuß von x über die größte ganze Zahl $\leqq x$ ist.

§ 4. Konvergente Folgen von Funktionen.

Es sei eine Folge von Funktionen $y_n = f_n(x)$ gegeben, sämtlich in derselben Menge A definiert, während alle Bildpunkte demselben metrischen Raume E_y angehören. Für jedes x sei die Folge der Bildpunkte konvergent und

(1) $$y = f(x) = \lim y_n = \lim f_n(x),$$

wodurch eine neue Funktion $y = f(x)$ in A definiert ist. Wir betrachten gleichzeitig die Folge der reellen (nichtnegativen) Funktionen, nämlich der Entfernungen

(2) $$\eta_n = \varphi_n(x) = \overline{y y_n},$$

die in der ganzen Menge A nach $\lim \varphi_n(x) = 0$ konvergiert.

Hier erhebt sich vor allem die Frage, ob man aus der Stetigkeit aller Funktionen $f_n(x)$ auf die Stetigkeit der Limesfunktion $f(x)$ schließen darf. Bekannte einfache Beispiele, wie etwa $f_n(x) = \frac{1}{1+nx^2}$ für reelles x, wobei $f(x) = 0$ für $x \neq 0$ und $f(0) = 1$, zeigen, daß diese Frage zu verneinen ist. Wir werden aber sehen, daß unter gewissen Annahmen über A die Funktion $f(x)$ höchstens punktweise unstetig ist.

Wir erinnern zunächst an den Begriff der **gleichmäßigen** Konvergenz. Auf Grund der Konvergenz von η_n nach 0 existiert zu jedem positiven ε und jedem x eine natürliche Zahl ν_x derart, daß $\eta_n < \varepsilon$ für $n \geqq \nu_x$. Reicht man für alle x mit demselben $\nu_x = \nu$ aus, ist also

(3) $$\eta_n < \varepsilon \text{ für } n \geqq \nu \text{ und alle } x,$$

und läßt sich für jedes ε ein solches ν finden, so heißt die Funktionenfolge in A gleichmäßig konvergent.

Weniger verlangt folgende Bedingung, die wir in Ermanglung eines allgemein akzeptierten Namens als **uniforme** Konvergenz bezeichnen wollen. Die Funktionenfolge[1] heiße in A uniform kon-

[1] Sie wird hier und im folgenden stets in A konvergent angenommen; an sich ist z. B. die Bedingung (4) schon erfüllt, wenn $f(x) = f_1(x)$, gleichviel wie sich die übrigen Funktionen verhalten. Der obige Sprachgebrauch kollidiert übrigens mit dem französischen usw., wo uniform = gleichmäßig ist.

vergent, wenn für jedes ε wenigstens eine natürliche Zahl ν derart existiert, daß

(4) $$\eta_\nu < \varepsilon \text{ für alle } x.$$

Es ist evident, daß die gleichmäßige Konvergenz ein spezieller Fall der uniformen ist.

Wir übertragen diese Begriffe von der ganzen Menge A auf ihre einzelnen Punkte. Die Funktionenfolge heißt im Punkte a gleichmäßig konvergent (oder a ein Punkt gleichmäßiger Konvergenz), wenn für jedes positive ε eine Umgebung U_a und eine natürliche Zahl ν derart existiert, daß

(5) $$\eta_n < \varepsilon \text{ für } n \geqq \nu \text{ und alle } x \text{ in } U_a;$$

sie heißt in a uniform konvergent (a ein Punkt uniformer Konvergenz), wenn für jedes positive ε eine Umgebung U_a und eine natürliche Zahl ν derart existiert, daß

(6) $$\eta_\nu < \varepsilon \text{ für alle } x \text{ in } U_a.$$

Die Punkte gleichmäßiger Konvergenz sind zugleich Punkte uniformer Konvergenz. Aus diesen Definitionen folgt nun genau wie für die Menge der Stetigkeitspunkte einer Funktion, daß auf den Raum A bezogen sowohl die Menge der Punkte gleichmäßiger wie die der Punkte uniformer Konvergenz von der Form G_δ ist. Bezeichnen wir nämlich die Menge der Punkte a, für die bei festem ε die Bedingung (5) erfüllbar ist, mit $A(\varepsilon)$, so ist wie in § 3

$$\mathfrak{D}\left(A(1), A(\tfrac{1}{2}), A(\tfrac{1}{3}), \ldots\right)$$

die Menge der Punkte gleichmäßiger Konvergenz, und wie dort ist leicht ersichtlich, daß jedes x von U_a selbst ein Punkt von $A(\varepsilon)$, a also ein innerer Punkt dieser Menge und $A(\varepsilon)$ daher ein Gebiet ist. Das Gleiche gilt von der uniformen Konvergenz. Wir bemerken also:

I. Die Menge der Punkte gleichmäßiger und die der Punkte uniformer Konvergenz sind Durchschnitte von Gebietsfolgen, die der Punkte ungleichmäßiger und die der Punkte nichtuniformer Konvergenz sind Summen von Folgen abgeschlossener Mengen.

Hier wie bei der Trennung von Stetigkeits- und Unstetigkeitspunkten spaltet sich also A in zwei Borelsche Mengen (S. 305): $A = G_\delta + F_\sigma$. Wenn die Funktionenfolge in A gleichmäßig konvergiert, so sind alle Punkte von A Punkte gleichmäßiger Konvergenz. Das Umgekehrte ist nicht richtig; man darf nicht einmal schließen, daß zu einem Punkt a gleichmäßiger Konvergenz eine

Umgebung U_a gehöre, in der $f_n(x)$ gleichmäßig konvergiert, denn natürlich kann in der Formel (5) bei abnehmendem ε auch die zugehörige Umgebung U_a unbegrenzt abnehmen so, daß der Durchschnitt aller U_a nur aus dem Punkte a besteht. Das Gleiche gilt für uniforme Konvergenz.

Wenn indessen die Menge A in sich kompakt ist und eine abzählbare dichte Teilmenge hat, so läßt sich auf Grund des Borelschen Satzes behaupten, daß die Funktionenfolge in A gleichmäßig konvergiert, falls sie in jedem Punkte von A gleichmäßig konvergiert. Denn von den Umgebungen U_a, die bei gegebenem ε die Ungleichung (5) erfüllen, schließt eine endliche Summe die Menge A ein; wählt man von den zugehörigen endlich vielen Zahlen ν die größte und bezeichnet sie mit ν, so gilt die Ungleichung (3), also konvergiert die Funktionenfolge in A gleichmäßig. Für uniforme Konvergenz läßt sich der entsprechende Schluß nicht ziehen, sondern nur behaupten, daß (falls jeder Punkt von A ein Punkt uniformer Konvergenz ist) die Ungleichung $\eta_\nu < \varepsilon$ bei vorgeschriebenem ε für alle x durch eine endliche Menge von Zahlen ν bewirkt wird.

Nach diesen Vorbereitungen können wir die Frage, ob eine Folge stetiger Funktionen nach einer wieder stetigen Funktion konvergiert, durch den folgenden, im Prinzip von C. Arzelà herrührenden Satz beantworten:

II. Sind alle $f_n(x)$ in a stetig, so ist $f(x)$ in a stetig dann und nur dann, wenn a ein Punkt uniformer Konvergenz ist.

Beweis. Wir setzen wie immer

$$b = f(a), \; y = f(x), \; b_n = f_n(a), \; y_n = f_n(x).$$

Ist nun a ein Punkt uniformer Konvergenz, so gibt es zu vorgeschriebenem ε ein ν und ein U_a derart, daß, für jedes x in U_a, η_ν oder $\overline{yy_\nu} < \varepsilon$, insbesondere auch $\overline{bb_\nu} < \varepsilon$. Wegen der Stetigkeit von $f_\nu(x)$ in a kann man, indem man eventuell U_a verkleinert, annehmen, daß auch $\overline{b_\nu y_\nu} < \varepsilon$. Aus diesen drei Ungleichungen folgt $\overline{by} < 3\varepsilon$, wenn x in U_a liegt, also ist $f(x)$ in a stetig.

Umgekehrt, ist $f(x)$ in a stetig, so bestimme man auf Grund der Konvergenz zunächst ein ν so, daß $\overline{bb_\nu} < \varepsilon$, sodann, auf Grund der Stetigkeit von $f(x)$ und $f_\nu(x)$, eine Umgebung U_a, für deren Punkte $\overline{by} < \varepsilon$ und $\overline{b_\nu y_\nu} < \varepsilon$. Aus diesen drei Ungleichungen folgt $\overline{yy_\nu} < 3\varepsilon$, wenn x in U_a liegt, d. h. die uniforme Konvergenz in a.

Sind die $f_n(x)$ in der ganzen Menge A stetig, so ist also die Limesfunktion $f(x)$ in den Punkten uniformer Konvergenz und nur

in diesen stetig; wir machen auf die Bestätigung des Satzes von § 3 durch den gegenwärtigen Satz I aufmerksam, hinsichtlich der Form G_δ für die Menge der Stetigkeitspunkte wie für die Menge der Punkte uniformer Konvergenz. Spezielle Fälle des Satzes II, die aber nur hinreichende, nicht notwendige Kriterien liefern, sind:

III. Sind die $f_n(x)$ in A stetig, so ist $f(x)$ ebenfalls in A stetig, wenn zu jedem Punkt a eine Umgebung U_a gehört, in der die Funktionenfolge uniform (im speziellen gleichmäßig) konvergiert; oder wenn die Funktionenfolge in A uniform (im speziellen gleichmäßig) konvergiert.

Diese Kriterien, meistens das allerspeziellste der gleichmäßigen Konvergenz in der ganzen Menge A, dienen in der Regel zur Feststellung der Stetigkeit einer durch eine konvergente Folge stetiger Funktionen definierten Limesfunktion. Anwendungen davon sind dem Leser aus den Elementen der Differentialrechnung und Funktionentheorie geläufig.

Wir nehmen jetzt an, daß die Menge A, in der die Funktionen $f_n(x)$ definiert sind, eine Menge G_δ in einem vollständigen Raume E_x sei,[1] also nach Kap. VIII, § 9 am Ende eine Menge zweiter Kategorie in bezug auf sich selbst, d. h. A ist nicht die Summe einer Folge von Mengen, die in A nirgendsdicht sind. Wir wollen dann zeigen, daß, falls alle $f_n(x)$ in A stetig sind, die Limesfunktion $f(x)$ höchstens punktweise unstetig ist, oder daß die Menge der Punkte uniformer Konvergenz in A dicht ist. Wir zeigen sogar, daß schon die Menge der Punkte gleichmäßiger Konvergenz in A dicht ist.

Es sei, bei einem bestimmten positiven ε, A_ν die Menge der Punkte x von A, in denen $\eta_n < \varepsilon$ für $n \geqq \nu$. Es ist $A_1 \leqq A_2 \leqq A_3 \leqq \ldots$, und jedes x gehört einmal einer Menge A_ν (und allen folgenden) an, also

$$A = \mathfrak{S}(A_1, A_2, A_3, \ldots).$$

Nach der Voraussetzung ist es ausgeschlossen, daß alle Mengen A_ν in A nirgendsdicht seien; es sei also A_ν nicht in A nirgendsdicht, d. h. A_ν zu einem von Null verschiedenen Relativgebiet A' von A dicht, sodaß in jeder Umgebung U_a eines Punktes a von A' sich Punkte von A_ν befinden.

Ist a ein Punkt von A', so gibt es wegen der Konvergenz der Funktionenfolge eine natürliche Zahl $p \geqq \nu$ so, daß $\overline{b b_p} < \varepsilon$. Ist ferner n eine beliebige natürliche Zahl $\geqq \nu$, so läßt sich wegen der Stetigkeit von $f_p(x)$ und $f_n(x)$ eine Umgebung U_a angeben, für

[1] Aber die Umgebungen U_a, von denen weiterhin die Rede ist, beziehen sich wie immer auf A (nicht E_x) als umfassenden Raum.

deren Punkte $\overline{b_p y_p} < \varepsilon$ und $\overline{b_n y_n} < \varepsilon$. Insbesondere wählen wir für x einen in U_a liegenden Punkt von A_ν, dann ist zugleich $\overline{y y_p} < \varepsilon$ und $\overline{y y_n} < \varepsilon$. Aus diesen fünf Ungleichungen folgt $\overline{b b_n} < 5\varepsilon$.

Vertauschen wir ε mit $\frac{1}{5}\varepsilon$, so läßt sich also zu gegebenem ε eine natürliche Zahl ν und ein nichtverschwindendes Relativgebiet A' von A angeben, für dessen sämtliche Punkte $\overline{b b_n} < \varepsilon$ ist, sobald $n \geqq \nu$. Da auch alle Punkte einer hinlänglich kleinen Umgebung U_a (nämlich für $U_a \leqq A'$) dieselbe Eigenschaft haben, so ist a ein Punkt der Menge $A(\varepsilon)$, für deren Punkte die Bedingung (5) erfüllbar ist. Diese Menge ist also jedenfalls, bei beliebigem ε, von Null verschieden.

Wenden wir dasselbe Raisonnement, statt auf A, auf irgend ein nichtverschwindendes Relativgebiet von A an, das ja wieder eine Menge G_δ in E_x ist, so folgt, daß auch dieses mindestens einen Punkt von $A(\varepsilon)$ enthält, d. h. die Menge $A(\varepsilon)$ ist in A dicht. Nach Kap. VIII, § 9, VII ist dann auch der Durchschnitt

$$\mathfrak{D}\left(A(1),\ A(\tfrac{1}{2}),\ A(\tfrac{1}{3}),\ \ldots\right),$$

d. h. die Menge der Punkte gleichmäßiger Konvergenz, immer noch in A dicht. Damit ist unser Ziel erreicht und der Satz bewiesen:

IV. **Ist A eine Menge G_δ in einem vollständigen Raume, und konvergiert die Folge der in A stetigen Funktionen $f_n(x)$ nach $f(x)$, so ist die Menge der Punkte gleichmäßiger Konvergenz in A dicht. Um so mehr ist also die Menge der Punkte uniformer Konvergenz in A dicht, d. h. die Funktion $f(x)$ höchstens punktweise unstetig.**

Danach ist eine in A überall unstetige Funktion, z. B. die **Dirichletsche Funktion** einer reellen Variablen, die für rationales x gleich Eins und für irrationales gleich Null ist, nicht durch eine konvergente Folge stetiger Funktionen darstellbar.

Ohne die Voraussetzung, daß A ein G_δ eines vollständigen Raumes sei, braucht der Satz nicht zuzutreffen. Es sei z. B. $A = \{a_1, a_2, \ldots\}$ die Menge der rationalen Zahlen, also kein solches G_δ (S. 321). Jede beliebige (reelle) Funktion $f(x)$ läßt sich hier als Limes einer Folge stetiger Funktionen darstellen, denn ist $f(a_p) = b_p$ beliebig vorgeschrieben, so kann man immer eine stetige Funktion $f_n(x)$, z. B. ein Polynom, wählen, die für $x = a_1, a_2, \ldots, a_n$ die Werte $b_1, b_2, \ldots, b_n$ annimmt. Es ist dann $f_n(a_p) = b_p$ für $n \geqq p$, also $\lim f_n(a_p) = b_p$. Dabei kann $f(x)$ in A überall unstetig gemacht werden, z. B. indem man an den dyadisch rationalen Stellen $f(x) = 0$, an den übrigen $f(x) = 1$ vorschreibt.

Für die Darstellbarkeit von $f(x)$ als Limes einer Folge stetiger

Funktionen ist nach IV (unter der über A gemachten Voraussetzung) notwendig, daß $f(x)$ höchstens punktweise unstetig sei. Daß diese Bedingung keinesfalls hinreicht, geht schon aus einer einfachen Mächtigkeitsbetrachtung hervor. Wenn A unendlich und der Raum $E_x \geqq A$ sowie der Bildraum E_y, dem die Funktionswerte angehören, euklidisch ist, so ist die Menge der in A stetigen Funktionen von der Mächtigkeit $\aleph$ des Kontinuums (S. 365); es gibt dann auch nur $\aleph^{\aleph_0} = \aleph$ Folgen stetiger Funktionen und $\aleph$ Limesfunktionen konvergenter Folgen stetiger Funktionen. Andererseits gibt es aber, bei geeignetem A, $2^{\aleph} > \aleph$ punktweise unstetige Funktionen. Sei z. B. $A = E_x$, F eine nirgendsdichte, abgeschlossene Menge von der Mächtigkeit $\aleph$ und G das komplementäre, dichte Gebiet;[1] definieren wir dann eine Funktion $f(x)$, die in G Null und in F von Null verschieden ist, so ist $f(x)$ in den Punkten von G stetig, in denen von F unstetig (letzteres, weil jeder Punkt von F Häufungspunkt von G ist). Die Funktion ist also punktweise unstetig, und da ihre Werte in F beliebig sind, so gibt es $2^{\aleph}$ solcher Funktionen.

Wir können auch die notwendigen Bedingungen dafür, daß $f(x)$ Limes stetiger Funktionen sei, leicht verschärfen. Ist P in A abgeschlossen, also ebenfalls von der Form G_δ, und beschränkt man die Variabilität von x auf P, so darf diese Teilfunktion $f(x|P)$ offenbar in P ebenfalls nur punktweise unstetig sein, und das muß für jede solche Menge P gelten. Daß diese Bedingung dann auch hinreichend ist, hat — allerdings unter weniger allgemeinen Voraussetzungen als den unserigen — R. Baire bewiesen.

Von Baire stammt auch die nachstehende, den Cantorschen Ordnungszahlen folgende Klassifikation der durch Limesbildung entstehenden Funktionen. Denken wir uns ein System Φ von Funktionen $y = f(x)$, sämtlich in derselben Menge A definiert, während y Punkt eines Raumes E_y ist. Dies System sei in dem Sinne abgeschlossen, daß der Limes $f(x) = \lim f_n(x)$ einer konvergenten Folge von Funktionen des Systems wieder dem System angehört, und zwar sei Φ das kleinste abgeschlossene System über dem System Φ_0 der stetigen Funktionen. (Es gibt abgeschlossene Systeme, z. B. das aller Funktionen, und der Durchschnitt aller abgeschlossenen Systeme $\geqq \Phi_0$ ist das System Φ.) Dieses System ist nun, in derselben Art wie ein $(\sigma\delta)$-System von Mengen (S. 305), folgendermaßen durch transfinite Induktion aufzubauen: Φ_1 bestehe aus den Limesfunktionen von Folgen stetiger Funktionen, Φ_2 aus den Limesfunktionen von Funktionen des Systems Φ_1 usw.; allgemein bestehe,

[1] Z. B. A die Menge der reellen Zahlen und F die Menge der Zahlen, deren triadische Entwicklung nur Nullen und Zweien aufweist (vgl. S. 254).

für $0 < \eta < \omega_1$, Φ_η aus den Limesfunktionen konvergenter Folgen von Funktionen des Systems $\underset{\xi}{\mathfrak{S}} \Phi_\xi$ $(\xi < \eta)$. Dann ist $\Phi = \underset{\eta}{\mathfrak{S}} \Phi_\eta$, wobei indessen nicht ausgeschlossen ist, daß der Prozeß schon an einer früheren Stelle abbricht und, für ein bestimmtes η, bereits $\Phi_\eta = \Phi_{\eta+1} = \ldots = \Phi$ ist. Man bemerke übrigens, daß unter denselben Annahmen wie vorhin das System Φ nur von der Mächtigkeit des Kontinuums ist, also nur einen winzigen Bruchteil von der Menge aller unstetigen Funktionen umfaßt.

Z. B. gehört die Ableitung einer differenzierbaren Funktion $f(x)$ zum System Φ_1, denn $f'(x) = \lim n\left[f\left(x + \frac{1}{n}\right) - f(x)\right]$ ist Limes stetiger Funktionen. Die Dirichletsche Funktion gehört zum System $\Phi_2 - \Phi_1$ (S. 267).

Die Baireschen Funktionen, wie wir die Funktionen des Systems Φ nennen wollen, haben, falls A wieder eine Menge G_δ in einem vollständigen Raume ist, die gemeinsame Eigenschaft[1]: für jede existiert eine in A dichte Menge P der Form G_δ derart, daß die auf P beschränkte Teilfunktion $f(x|P)$ stetig ist. Da dies für die stetigen Funktionen offenbar zutrifft (mit $P = A$), so brauchen wir nur zu zeigen, daß die genannte Eigenschaft bei Limesbildung erhalten bleibt. Ist $f(x) = \lim f_n(x)$ und $f_n(x|P_n)$ stetig, wo P_n von der Form G_δ und in A dicht ist, so ist der Durchschnitt $D = \mathfrak{D}(P_1, P_2, \ldots)$ ebenfalls ein G_δ und in A dicht (S. 326). Ferner sind alle Funktionen $f_n(x|D)$ stetig, also $f(x|D)$ höchstens punktweise unstetig, und die Menge P der Stetigkeitspunkte dieser Funktion ist ein in D dichtes G_δ. Da P in D, D in A dicht ist, so ist auch P in A dicht und $f(x|P)$ stetig. — Z. B. ist die Dirichletsche, überall unstetige Funktion bei Beschränkung auf die irrationalen Stellen, die ja eine dichte Menge G_δ bilden, konstant $= 0$, also stetig.

§ 5. Funktionenklassen.

Die Betrachtungen über konvergente Folgen lassen sich, zunächst für reelle Funktionen, unmittelbar mit denen in Kap. I, § 11 in Verbindung bringen. Wir nannten dort eine in A definierte reelle Funktion $f(x)$ von der Klasse (M, N), wenn, für jedes reelle u, die Menge der Punkte x, wo $f(x) > u$, eine Menge M, und die Menge der Punkte x, wo $f(x) \geqq u$, eine Menge N ist; dabei dachten wir uns M und N gewisse Mengensysteme durchlaufend. Unter der auch im folgenden festzuhaltenden Annahme, daß die M einen σ-Ring,

[1] Weitere Eigenschaften ähnlichen Charakters beruhen auf der Lebesgueschen Maßtheorie (Kap. X, § 4).

die N einen δ-Ring bilden und daß sie Komplemente von einander sind ($A-M$ ist ein N, $A-N$ ein M), hatten wir gefunden, daß der Limes einer konvergenten Folge von Funktionen der Klasse (M, N) seinerseits von der Klasse (N_σ, M_δ) ist. Da die Mengen N_σ, M_δ die über M, N gemachten Voraussetzungen ebenfalls erfüllen, so ist der Limes einer konvergenten Folge von Funktionen der Klasse (N_σ, M_δ) eine Funktion der Klasse $(M_{\delta\sigma}, N_{\sigma\delta})$ usw. Nun sind nach dem Satze § 1, III die stetigen Funktionen von der Klasse (G, F), wo G ein Gebiet, F eine abgeschlossene Menge (beides bezogen auf den Raum A) bedeutet; denn in der Menge B der reellen Zahlen $f(x)$ ist die Menge B' derer, die $>u$ sind, ein Relativgebiet, also ihr Urbild A', d. h. die Menge der Punkte x von A, wo $f(x)>u$, ein Relativgebiet in A, und das Entsprechende gilt bezüglich der Ungleichung $f(x) \geqq u$. Die Mengen G, F erfüllen die obigen Voraussetzungen über M, N; demnach ist eine Limesfunktion stetiger Funktionen, d. h. eine Bairesche Funktion des Systems Φ_1, von der Klasse (F_σ, G_δ), eine Funktion des Systems Φ_2 von der Klasse $(G_{\delta\sigma}, F_{\sigma\delta})$ und so fort: den Baireschen Funktionen entsprechen die Borelschen Mengen. Es wäre von Interesse, zu untersuchen, ob diese Kriterien umkehrbar sind (daß eine Funktion der Klasse (G, F) stetig ist, werden wir sogleich sehen).

Wir formen den Klassenbegriff etwas um, so daß er sich in dieser neuen Gestalt auf beliebige Funktionen übertragen läßt. Eine reelle Funktion ist dann und nur dann von der Klasse (M, N), wenn jedem Relativgebiet der Bildmenge B eine Menge M, jeder in B abgeschlossenen Menge eine Menge N als Urbild entspricht. Denn wenn $f(x)$ von der Klasse (M, N) ist, so ist die Menge der x, wo

$$f(x)>u, \quad f(x)<v, \quad u<f(x)<v,$$

resp. ein M, ein $A-N=M$, ein $\mathfrak{D}(M, M)=M$; und da jedes lineare Gebiet die Summe von offenen Strecken, eventuell Halbgeraden, in höchstens abzählbarer Menge ist, so ist das Urbild jedes Relativgebietes von B eine Menge M oder $M_\sigma = M$; das Urbild jeder in B abgeschlossenen Menge eine Menge $A-M=N$. Damit ist die Behauptung bewiesen; insbesondere folgt nach § 1, III, daß eine Funktion der Klasse (G, F) stetig ist.

Demgemäß nennen wir eine beliebige Funktion $y=f(x)$ von der Klasse (M, N), wenn jedem Relativgebiet der Bildmenge B eine Menge M, jeder in B abgeschlossenen Menge eine Menge N als Urbild entspricht. Die Funktionen der Klasse (G, F) sind die stetigen Funktionen. Um das dem früheren analoge Resultat über den Limes einer konvergenten Folge abzuleiten, nehmen wir an, daß alle Funktionswerte oder Bildpunkte einem metrischen Raum E_y mit (höchstens) abzählbarer dichter Teilmenge angehören.

Ist a ein fester, x ein variabler Punkt von A, $f(x)$ eine Funktion der Klasse (M, N), so ist die als Funktion von x betrachtete Entfernung der Bildpunkte $\varphi(x) = \overline{by}$ offenbar von derselben Klasse; denn die Ungleichung $\varphi(x) > u$ definiert in B ein Relativgebiet, also in A eine Menge M (analog für $\varphi(x) \geqq u$). Umgekehrt aber: wenn die Funktion $\varphi(x)$ stets, für jeden Punkt a, von der Klasse (M, N) ist, so ist auch $f(x)$ von dieser Klasse; denn da die Ungleichung $\varphi(x) < u$ mit positivem u in A eine Menge M, in B eine Umgebung V_b definiert, so hat jedes V_b ein Urbild M, jedes Relativgebiet von B, als Summe höchstens abzählbar vieler Umgebungen, ein Urbild M oder $M_\sigma = M$. Für den Limes $f(x) = \lim f_n(x)$ einer konvergenten Folge von Funktionen der Klasse (M, N) folgt demgemäß, wenn man die zugehörigen Funktionen

$$\varphi(x) = \overline{by}, \quad \varphi_n(x) = \overline{b_n y_n}, \quad \varphi(x) = \lim \varphi_n(x)$$

betrachtet, daß $f(x)$ von der Klasse (N_σ, M_δ) ist; die Baireschen Funktionen sind also auch hier von den Klassen (G, F), (F_σ, G_δ), $(G_{\delta\sigma}, F_{\sigma\delta})$, ...

Wir fügen noch folgende Betrachtung hinzu (das Analogon derjenigen über die Summe zweier Funktionen, S. 29). Sind $y_1 = f_1(x)$, $y_2 = f_2(x)$ zwei Funktionen der Klasse (M, N), so ist auch die als Funktion von x betrachtete Entfernung der Bildpunkte $\varphi(x) = \overline{y_1 y_2}$ von dieser Klasse, wieder vorausgesetzt, daß die Bildmengen B_1, B_2 einem metrischen Raume mit höchstens abzählbarer dichter Teilmenge angehören, oder daß sie höchstens abzählbare dichte Teilmengen C_1, C_2 haben. Wenn nämlich die Ungleichung $\overline{y_1 y_2} > u$ besteht, so läßt sich eine natürliche Zahl n und zwei Punkte c_1, c_2 aus C_1, C_2 so bestimmen, daß

$$\overline{c_1 y_1} < \frac{1}{n}, \quad \overline{c_2 y_2} < \frac{1}{n}, \quad \overline{c_1 c_2} > u + \frac{2}{n};$$

zu diesem Zwecke wähle man zuerst $\frac{4}{n} < \overline{y_1 y_2} - u$, dann c_1 und c_2 in Abständen $< \frac{1}{n}$ von y_1 und y_2, und dann ist

$$\overline{c_1 c_2} \geqq \overline{y_1 y_2} - \overline{c_1 y_1} - \overline{c_2 y_2} > u + \frac{4}{n} - \frac{1}{n} - \frac{1}{n}.$$

Umgekehrt ziehen die drei obigen Ungleichungen die Ungleichung $\overline{y_1 y_2} > u$ nach sich. Die Menge der Punkte x, wo $\varphi(x) > u$, wird also so gefunden: für irgend einen Komplex c_1, c_2, n nehme man die Menge der Punkte x, wo gleichzeitig $\overline{c_1 y_1} < \frac{1}{n}$, $\overline{c_2 y_2} < \frac{1}{n}$, das ist, da diese Ungleichungen Relativgebiete in B_1, B_2 definieren, der Durchschnitt zweier Mengen M oder selbst ein M; sodann bilde

man die Summe über alle Komplexe der Eigenschaft $\overline{c_1 c_2} > u + \frac{2}{n}$, das gibt höchstens abzählbar viele Summanden und demnach eine Menge $M_\sigma = M$. In derselben Weise zeigt man, daß der Ungleichung $\varphi(x) < u$ eine Menge M, also der Ungleichung $\varphi(x) \geqq u$ eine Menge $A - M = N$ entspricht: $\varphi(x)$ ist von der Klasse (M, N).

Sind speziell die Funktionen $f_1(x)$, $f_2(x)$ stetig, so ist weder das vorstehende Raisonnement noch die Voraussetzung (der höchstens abzählbaren dichten Teilmenge) über E_y nötig, um die Stetigkeit von $\varphi(x)$ zu beweisen. Zu jedem Punkte a und jedem $\varepsilon > 0$ gibt es eine Umgebung U_a, für deren Punkte $\overline{b_1 y_1} < \varepsilon$, $\overline{b_2 y_2} < \varepsilon$, also $|\overline{y_1 y_2} - \overline{b_1 b_2}| < 2\varepsilon$; $\varphi(x)$ ist in a stetig.

Wir fügen noch einige Betrachtungen über reelle Funktionen hinzu. Wenn die Menge der Stellen $f(x) > u$ stets ein M ist, so ist die Menge der Stellen $f(x) \geqq u$ jedenfalls ein M_δ (als Durchschnitt der Mengen $f(x) > u - 1, u - \frac{1}{2}, u - \frac{1}{3}, \ldots$); also in leicht verständlicher Schreibweise: eine Funktion der Klasse $(M, *)$ ist von der Klasse (M, M_δ), ebenso eine Funktion der Klasse $(*, N)$ von der Klasse (N_σ, N). Bilden die Funktionen $f_1(x), f_2(x), \ldots$ von der Klasse $(M, *)$ für jedes x eine nach oben beschränkte Zahlenfolge mit der oberen Schranke $f(x)$, so ist auch $f(x)$ von der Klasse $(M, *)$; denn die Menge $f(x) > u$ ist die Summe der Mengen $f_1(x) > u$, $f_2(x) > u, \ldots$. Dies gilt für beliebig (auch unabzählbar) viele Funktionen, falls die Summe beliebig vieler M wieder ein M ist, insbesondere für Gebiete; die obere Schranke beliebig vieler Funktionen der Klasse $(G, *)$ ist wieder von der Klasse $(G, *)$, die untere Schranke von Funktionen der Klasse $(*, F)$ wieder von der Klasse $(*, F)$.

Die Stetigkeitsforderung für den Punkt a besagte, daß zu jedem positiven ε eine Umgebung U_a existiert, in der

$$-\varepsilon < f(x) - f(a) < \varepsilon.$$

Die Funktionen, die nur die eine dieser beiden Ungleichungen erfüllen, werden nach R. Baire halbstetig genannt, und zwar heißt $f(x)$ an der Stelle a unterhalb stetig, wenn für jedes positive ε und geeignetes U_a

$$f(x) > f(a) - \varepsilon;$$

oberhalb stetig, wenn in demselben Sinne

$$f(x) < f(a) + \varepsilon.$$

Die in A unterhalb stetigen Funktionen sind mit den Funktionen der Klasse $(G, *)$, die oberhalb stetigen mit denen der Klasse $(*, F)$ identisch. Ist nämlich $f(x)$ unterhalb stetig und $f(a) > u$, so ist auch noch in einer gewissen Umgebung U_a $f(x) > u$, also a innerer Punkt der Menge $f(x) > u$ und diese Menge

ein Gebiet (wie immer auf den Raum A bezogen). Wenn umgekehrt $f(x)$ von der Klasse $(G, *)$ ist, so ist die Menge der x, wo $f(x) > f(a) - \varepsilon$, ein Gebiet, dem der Punkt a und also eine gewisse Umgebung U_a angehört. Das Entsprechende gilt für oberhalb stetige Funktionen. Danach ist die obere Schranke von unterhalb stetigen Funktionen wieder unterhalb stetig, die untere Schranke von oberhalb stetigen Funktionen wieder oberhalb stetig.

Das Verhalten einer Funktion $f(x)$ in der Nachbarschaft der Stelle a gibt noch zu folgenden Bestimmungen Anlaß. Nennen wir v eine reelle Zahl, die alle Werte $f(x)$ einer geeigneten Umgebung U_a übertrifft, u eine Zahl, die das nicht tut. Zahlen u gibt es jedenfalls, z. B. $f(a)$; wenn es auch Zahlen v gibt, d. h. wenn $f(x)$ in einer Umgebung U_a nach oben beschränkt ist, so ist $u < v$ und dieser Schnitt in der Zahlenmenge bestimmt eine Zahl $\varphi(a) \geqq f(a)$, die wir die obere Grenze von $f(x)$ im Punkte a nennen wollen.[1] Für jedes positive ε ist also ein U_a angebbar, in dem

$$f(x) < \varphi(a) + \varepsilon,$$

während andererseits in jedem U_a Punkte mit

$$f(x) > \varphi(a) - \varepsilon$$

vorhanden sind. Entsprechend ist die untere Grenze $\psi(a) \leqq f(a)$ von $f(x)$ im Punkte a zu erklären, falls $f(x)$ in einer Umgebung von a nach unten beschränkt ist.

Ist $f(x)$ in a oberhalb stetig, so ist $\varphi(a) = f(a)$ und vice versa; die Differenz $\varphi(a) - \psi(a)$, die man die Schwankung der Funktion im Punkte a nennt, ist Null oder positiv, je nachdem a ein Stetigkeits- oder Unstetigkeitspunkt von $f(x)$ ist.

Nehmen wir an, daß $\varphi(x)$ überall in A existiert, so ist dies eine oberhalb stetige Funktion. Denn in einem geeigneten U_a ist

$$f(x) < \varphi(a) + \varepsilon;$$

ist ferner $U_x \leqq U_a$, so gibt es in U_x einen Punkt y mit

$$f(y) > \varphi(x) - \varepsilon,$$

und da $f(y) < \varphi(a) + \varepsilon$, so folgt

$$\varphi(x) < \varphi(a) + 2\varepsilon.$$

Überdies ist $\varphi(x)$ die kleinste oberhalb stetige Funktion $\geqq f(x)$; denn soll $f'(x) \geqq f(x)$ oberhalb stetig sein, so ist $\varphi'(x) \geqq \varphi(x)$, d. h. $f'(x) \geqq \varphi(x)$. Man kann auf diesem Wege auch zur Definition von

[1] Wenn keine Zahlen v existieren, wird $\varphi(a)$ nicht definiert oder allenfalls $\varphi(a) = +\infty$ gesetzt. $\varphi(a)$ ist nicht mit dem unter Ausschluß der Stelle a erklärten $\limsup_{x=a} f(x)$ identisch (vgl. unten).

$\varphi(x)$ gelangen; denn wenn überhaupt oberhalb stetige Funktionen $\geqq f(x)$ existieren, so ist deren untere Schranke wieder oberhalb stetig und also die kleinste oberhalb stetige Funktion $\geqq f(x)$.

Ebenso ist $\psi(x)$ die größte unterhalb stetige Funktion $\leqq f(x)$. Die Differenz $\varphi(x) - \psi(x)$ ist oberhalb stetig, also von der Klasse (F_σ, F); die Menge der Stellen $\varphi(x) - \psi(x) > 0$, d. h. der Unstetigkeitspunkte von $f(x)$, ist also ein F_σ, die der Stetigkeitspunkte ein G_δ, wie wir schon wissen.

Dieselben Betrachtungen lassen sich auch für einen Punkt a anstellen, der gar nicht dem Argument A der Funktion $f(x)$ angehört, wohl aber Häufungspunkt von A in einem Raume $E_x \geqq A$ ist. Die Zahlen $\varphi(a)$, $\psi(a)$ werden hier als

$$\limsup_{x=a} f(x), \quad \liminf_{x=a} f(x)$$

bezeichnet, im Falle der Gleichheit als $\lim_{x=a} f(x)$, oder, um keinen Zweifel zu lassen:

Es sei U_a der Durchschnitt einer Umgebung von a (im Raume E_x) mit der Menge A; $U_a \leqq A$ enthält unendlich viele Punkte x, die sämtlich von a verschieden sind. v sei eine reelle Zahl, die alle Werte $f(x)$ einer geeigneten Menge U_a übertrifft, u eine Zahl, die das nicht tut. Wenn es sowohl Zahlen v als auch Zahlen u gibt,[1] so ist $u < v$ und der Schnitt bestimmt definitionsgemäß die Zahl $\limsup_{x=a} f(x)$.

Endlich kann man, wenn a der Menge A angehört und zugleich Häufungspunkt ist, einerseits mit Einschluß von a die Zahlen $\varphi(a)$, $\psi(a)$, andererseits mit Ausschluß von a, also für die Teilfunktion $f(x\ A - \{a\})$, die Zahlen $\limsup_{x=a} f(x)$, $\liminf_{x=a} f(x)$ in Betracht ziehen, wobei offenbar

$$\varphi(a) = \max f(a), \limsup f(x),$$
$$\psi(a) = \min f(a), \liminf f(x).$$

Die Funktion ist in a oberhalb stetig für $f(a) \geqq \limsup f(x)$, unterhalb stetig für $f(a) \leqq \liminf f(x)$, stetig für $f(a) = \lim f(x)$.

Wir wollen noch die Klasse der Ableitungen einer im abgeschlossenen Intervall (a, b) stetigen Funktion $f(x)$ bestimmen; wir denken uns die Funktion links von a mit $f(x) = f(a)$, rechts von b mit $f(x) = f(b)$ fortgesetzt und lassen x alle reellen Zahlen durchlaufen. Wenn $f(x)$ differenzierbar ist, so hatten wir schon gesehen

[1] Hier ist auch die Existenz von Zahlen u nicht sicher. Wenn es kein v resp. kein u gibt, pflegt man $\limsup_{x=a} f(x) = +\infty$ resp. $-\infty$ zu setzen.

(S. 390), daß $f'(x)$ als Limes stetiger Funktionen von der Klasse (F_σ, G_δ) ist. Bezeichnen wir die Differenzenquotienten mit

$$f(x, \xi) = \frac{f(\xi) - f(x)}{\xi - x}$$

und verstehen unter h eine positive Variable, so werden die Zahlen

$$\limsup_{h=0} f(x, x+h), \quad \liminf_{h=0} f(x, x+h),$$

falls sie existieren, als obere und untere rechte Ableitung von $f(x)$ an der Stelle x bezeichnet, ebenso die Zahlen

$$\limsup_{h=0} f(x, x-h), \quad \liminf_{h=0} f(x, x-h)$$

als obere und untere linke Ableitung. Stimmen die beiden rechten (linken) Ableitungen überein, so nennt man ihren gemeinsamen Wert die rechte (linke) Ableitung; stimmen alle vier überein, so erhält man die Ableitung schlechthin.

Betrachten wir etwa die obere rechte Ableitung

$$g(x) = \limsup_{h=0} f(x, x+h);$$

wir nehmen an, daß sie überall existiert. Bei festem x ist also, für ein hinlänglich kleines (von x abhängiges) k, $f(x, x+h)$ für $h < k$ nach oben beschränkt; da aber $f(x)$ beschränkt, $|f(x)| \leqq \mu$ ist, so ist auch für $h \geqq k$ $|f(x, x+h)| \leqq \frac{2\mu}{k}$ beschränkt, also $f(x, x+h)$ bei festem x für alle h nach oben beschränkt. Sei hiernach k konstant, $F(x, k)$ die obere Schranke von $f(x, x+h)$ für $h < k$, also, da bei festem h $f(x, x+h)$ stetige Funktion von x ist, $F(x, k)$ eine unterhalb stetige Funktion von x, von der Klasse (G, G_δ). Schließlich ist, wie leicht zu sehen, $g(x)$ die untere Schranke aller $F(x, k)$ für $k > 0$ oder, da diese Funktion mit k monoton abnimmt,

$$g(x) = \lim_n F\left(x, \frac{1}{n}\right)$$

von der Klasse $(G_{\delta\sigma}, G_\delta)$.

Die beiden oberen Ableitungen sind also von der Klasse $(G_{\delta\sigma}, G_\delta)$, die beiden unteren von der Klasse $(F_\sigma, F_{\sigma\delta})$. Wenn eine rechte (linke) oder eine einzige Ableitung existiert, so ist sie von der Klasse (F_σ, G_δ). Dies gilt zunächst absolut, bezogen auf die Menge aller reellen Zahlen, und dann natürlich relativ, bezogen auf das Intervall (a, b).

§ 6. Die Konvergenzpunkte einer Funktionenfolge.

Wenn die Folge der Funktionen $f_n(x)$ in A nicht überall konvergiert, so sei C die Menge der Punkte, wo sie konvergiert, D die Menge derer, wo sie nicht konvergiert, also $A = C + D$. Über diese

Mengen läßt sich im allgemeinen natürlich nichts behaupten, da man ja jedem Punkt x eine beliebige Punktfolge $f_n(x)$ zuordnen kann. Wenn indessen die Funktionen $f_n(x)$ stetige oder Bairesche Funktionen sind, so ist schon aus Mächtigkeitsgründen eine spezielle Beschaffenheit der Mengen C, D vorauszusagen.

Für eine Folge reeller Funktionen von der Klasse (M, N) wissen wir aus Kap. I, § 11, unter Festhaltung der bekannten Voraussetzungen über die Mengen M und N, daß C ein $N_{\sigma\delta}$ und D ein $M_{\delta\sigma}$ ist. Für eine Folge reeller, in A stetiger Funktionen bilden also die Konvergenzpunkte eine Menge $F_{\sigma\delta}$, die übrigen eine Menge $G_{\delta\sigma}$, bezogen auf den Raum A.

Die naheliegende Frage, ob sich nicht dieses Resultat noch vereinfachen läßt, ist zu verneinen, wie folgende Beispiele von Funktionen einer reellen Variablen lehren.

(α) Die Folge

$$y_1 = \sin 2\pi x,\ y_2 = \sin 4\pi x,\ y_3 = \sin 8\pi x,\ \ldots,\ y_n = \sin 2^n \pi x,\ \ldots$$

konvergiert für alle dyadisch rationalen Zahlen x, indem dann von einem gewissen Index ab $y_n = 0$ wird. Wir zeigen, daß sie für kein anderes x konvergiert. Wegen

$$y_{n+1}{}^2 = 4y_n{}^2(1-y_n{}^2), \quad y_{n+2} = 2y_{n+1}(1-2y_n{}^2)$$

muß im Falle der Existenz von $y = \lim y_n$

$$3y^2 - 4y^4 = 0, \quad y - 4y^3 = 0,$$

also $y = 0$ sein. Ist nun x nicht dyadisch rational, so kommt in seiner dyadischen Entwicklung

$$x = x_0 + \frac{x_1}{2} + \frac{x_2}{2^2} + \frac{x_3}{2^3} + \cdots$$

jede der beiden Ziffern 0, 1 unendlich oft vor; also ist, für unendlich viele n, gleichzeitig $x_{n+1} = 0$ und $x_{n+2} = 1$, demnach

$$2^n x = k + \frac{0}{2} + \frac{1}{2^2} + \frac{x_{n+3}}{2^3} + \ldots = k + \vartheta,$$

wo k eine ganze Zahl und $\frac{1}{4} < \vartheta < \frac{1}{2}$ ist. Daraus folgt

$$\sin 2^n \pi x = (-1)^k \sin \pi\vartheta, \ |y_n| = \sin \pi\vartheta > \frac{1}{\sqrt{2}},$$

und da diese Ungleichung unendlich oft erfüllt ist, so konvergiert y_n nicht nach 0, also überhaupt nicht.

Hier ist also C, die Menge der Konvergenzstellen, die Menge der dyadisch rationalen Zahlen, d. h. eine abzählbare insichdichte Menge, folglich keine Menge G_δ, und D ist kein F_σ.

(β) Mit Benutzung des vorigen Beispiels sei

$$z_n = (-1)^n : (1 + y_1{}^2 + y_2{}^2 + \cdots + y_n{}^2).$$

Ist x dyadisch rational, so wird für einen gewissen Index

$$z_n = -z_{n+1} = z_{n+2} = -z_{n+3} = \dots \neq 0,$$

die Folge konvergiert nicht. Ist x nicht dyadisch rational, so folgt aus dem unendlich häufigen Bestehen von $y_n^2 > \frac{1}{2}$, daß $\lim z_n = 0$. Hier haben also C und D gegen vorhin die Rollen getauscht; C ist kein F_σ und D kein G_δ.

Die Menge der Konvergenzstellen braucht also keine Menge F_σ oder G_δ zu sein (natürlich erst recht keine abgeschlossene Menge F und kein Gebiet G), und die Tatsache, daß sie stets ein $F_{\sigma\delta}$ ist, dürfte sich kaum vereinfachen lassen.

Für die Folgen von reellen Baireschen Funktionen des Systems Φ_1 ist die Konvergenzmenge ein $G_{\delta\sigma\delta}$ usw.

Um auch diese Ergebnisse von reellen Funktionen auf allgemeine zu übertragen, nehmen wir zunächst an, daß die Bildpunkte $y_n = f_n(x)$ einem vollständigen Raume E_y angehören. Unter dieser Voraussetzung erhalten wir folgende Darstellung der Konvergenzmenge C. Es sei $\varphi_{mn}(x) = \overline{y_m y_n}$; ferner, für positives ε, $A_{mn}(\varepsilon)$ die Menge derjenigen Punkte x, wo $\varphi_{mn}(x) \leqq \varepsilon$ ist;[1] sodann

$$A_m(\varepsilon) = \mathfrak{D}_n\, A_{mn}(\varepsilon) \text{ für } n = m+1,\ m+2,\ \dots$$

und

$$A(\varepsilon) = \mathfrak{S}\big(A_1(\varepsilon),\ A_2(\varepsilon),\ \dots\big).$$

Dann ist

$$C = \mathfrak{D}\big(A(1),\ A(\tfrac{1}{2}),\ A(\tfrac{1}{3}),\ \dots\big).$$

Dies ist nämlich in Formeln das Cauchysche Konvergenzkriterium, wonach die Folge y_n dann und nur dann konvergiert, wenn für jedes ε eine natürliche Zahl m angebbar ist derart, daß $\overline{y_m y_n} \leqq \varepsilon$ für $n > m$ (vgl. S. 314); oder wenn für jedes ε der Punkt x zu einer Menge $A_m(\varepsilon)$, also zu $A(\varepsilon)$ gehört. Demgemäß ist C der Durchschnitt aller $A(\varepsilon)$; da $A(\varepsilon) \leqq A(\varepsilon')$ für $\varepsilon < \varepsilon'$, so braucht wieder wie in früheren Fällen ε nicht alle positiven Zahlen, sondern nur eine nach Null konvergente Folge zu durchlaufen.

Sind nun die Funktionen $f_n(x)$ zunächst stetig, so ist (nach S. 393) jedes $\varphi_{mn}(x)$ stetig; dann ist $A_{mn}(\varepsilon)$ ein F, $A_m(\varepsilon)$ ein $F_\delta = F$, $A(\varepsilon)$ ein F_σ und C ein $F_{\sigma\delta}$. Der Satz, daß für eine Folge in A stetiger Funktionen die Konvergenzpunkte ein $F_{\sigma\delta}$, die übrigen ein $G_{\delta\sigma}$ bilden, gilt also auch, wenn die Funktionswerte einem vollständigen Raume angehören.

Hätten wir bei der Definition von $A_{mn}(\varepsilon)$ die Ungleichung $\varphi_{mn}(x) < \varepsilon$ statt $\leqq \varepsilon$ zugrunde gelegt, so hätte sich diese Menge

[1] Das Gleichheitszeichen nicht auszuschließen bringt hier eine Vereinfachung mit sich, wie wir sofort sehen werden.

als G, $A_m(\varepsilon)$ als G_δ, $A(\varepsilon)$ als $G_{\delta\sigma}$ und C als $G_{\delta\sigma\delta}$ ergeben: ein weniger präzises Resultat als das obige.

Sind die Funktionen $f_n(x)$ von der Klasse (M, N) und gehören ihre Werte einem vollständigen Raum mit höchstens abzählbarer dichter Teilmenge an, so folgt wiederum nach S. 392, daß die Funktionen $\varphi_{mn}(x)$ von der Klasse (M, N) sind; dann ist $A_{mn}(\varepsilon)$ ein N, $A_m(\varepsilon)$ ein $N_\delta = N$, $A(\varepsilon)$ ein N_σ und C ein $N_{\sigma\delta}$: die Konvergenzpunkte bilden eine Menge $N_{\sigma\delta}$, die übrigen eine Menge $M_{\delta\sigma}$, genau wie bei reellen Funktionen.

Zehntes Kapitel.

Inhalte von Punktmengen.

§ 1. Das Problem der Inhaltsbestimmung.

Schon in der Elementargeometrie und sodann in der Integralrechnung wird gewissen Punktmengen A des euklidischen Raumes E_3 ein *Volumen* $f_3(A)$ zugeschrieben, gewissen eine *Fläche* $f_2(A)$, gewissen eine *Länge* $f_1(A)$. Mengen mit nichtverschwindendem Volumen heißen *Körper* und haben keine (oder, wenn man will, „unendlich große“) Fläche und keine Länge; Mengen mit nichtverschwindender Fläche heißen *Flächen* und haben keine Länge, wohl aber verschwindendes Volumen; Mengen mit nichtverschwindender Länge heißen *Linien* oder Kurven und haben verschwindende Fläche und verschwindendes Volumen. Dies alles scheint bei Mengen elementaren Charakters sehr klar zu sein; aber schon eine Menge wie die der rationalen Punkte eines Würfels setzt uns in Verlegenheit, da sie weder mit Körpern noch Flächen noch Linien große Ähnlichkeit hat. Die Mengenlehre wird also den Versuch machen müssen, ob und in welchem Umfange man einer beliebigen Menge A solche Maßzahlen $f_1(A)$, $f_2(A)$, $f_3(A)$ zuordnen kann, die für gewöhnliche Punktmengen die gewöhnliche Bedeutung haben. Fügen wir noch $f_4(A) = f_5(A) = \ldots = 0$ hinzu und verallgemeinern das Problem in ersichtlicher Weise, so wird man für eine Punktmenge A_n im n-dimensionalen Raume ihren *m-dimensionalen Inhalt* $f_m(A_n)$ wenigstens von einem gewissen Index m ab zu definieren suchen, unter Zulassung oder Erwartung des Verschwindens von f_{m+1}, falls f_m definiert ist. Die wichtigsten und für die übrigen grundlegenden dieser Inhalte sind die n-dimensionalen Inhalte $f_n(A_n)$ von Mengen des n-dimensionalen Raumes, also die Längen linearer, die Flächen

ebener, die Volumina räumlicher Punktmengen usw.; wir werden sie ausschließlich in Betracht ziehen und schlechthin als Inhalte $f(A)$ bezeichnen.

Die fundamentale Wichtigkeit des Inhaltsbegriffs beruht vor allem auf seinem Zusammenhang mit dem Integralbegriff (§ 5): das Integral einer positiven Funktion wird einfach als Inhalt einer gewissen Punktmenge definiert.

Die Theorie des Inhalts hat sich in zwei Stufen entwickelt, die man am besten durch zwei Gesetze für additives Verhalten charakterisiert. Der Inhaltsbegriff der älteren Stufe, auf Ansätzen von G. Cantor und H. Hankel beruhend und von G. Peano und C. Jordan vervollständigt, erfüllt die Gleichung

$$f(A+B)=f(A)+f(B)$$

für zwei (oder endlich viele) paarweise fremde Mengen; der Inhaltsbegriff der neueren Stufe, den wir E. Borel und H. Lebesgue verdanken, erfüllt die Gleichung

$$f(A+B+C+\dots)=f(A)+f(B)+f(C)+\dots$$

für eine endliche oder abzählbare Menge von Summanden. Der älteren Stufe entspricht der Riemannsche, der neueren der Lebesguesche Integralbegriff. Der Übergang vom Endlichen zum Abzählbaren in der neueren Inhalts- und Integraltheorie darf als einer der größten Fortschritte der Mathematik bezeichnet werden, aus dem sehr einfachen Grunde, weil Folgen (von Zahlen, Funktionen, Punkten, Mengen) eben das Fundament der Analysis sind: jenes verschärfte Summengesetz macht den Inhaltsbegriff ohne weiteres von Mengenfolgen auf deren Summe und Durchschnitt, entsprechend den Integralbegriff von Funktionenfolgen auf deren Limes übertragbar.

Wir werden den Inhaltsbegriff konstruktiv behandeln, d. h. durch eine bestimmte Vorschrift die Zahlen $f(A)$ definieren und ihre Eigenschaften entwickeln. In der Wahl dieser Vorschrift läßt man sich natürlich durch gewisse Forderungen leiten, die an den Inhalt gestellt werden. Eine nur auf solchen Forderungen beruhende axiomatische Behandlungsweise ist von H. Lebesgue versucht worden, aber nicht zum Abschluß gekommen. Man wird jedenfalls das endliche, womöglich auch das abzählbare Summengesetz postulieren und überdies verlangen, daß kongruente[1] Mengen gleichen Inhalt haben; um endlich einen allen Inhalten gemeinsamen Proportionalitätsfaktor zu bestimmen und insbesondere das Verschwinden

[1] Dieser Begriff ist hier elementargeometrisch zu verstehen; es handelt sich um Mengen, die umkehrbar eindeutig und entfernungstreu aufeinander bezogen werden können.

aller Inhalte auszuschließen, wird man die Volumeneinheit fixieren, d. h. einem n-dimensionalen Würfel mit der Seitenlänge 1 den Inhalt 1 vorschreiben. Danach formuliert Lebesgue das Problem dahin, jeder (beschränkten) Menge A des Raumes E_n als Inhalt eine Zahl $f(A) \geqq 0$ unter folgenden Bedingungen zuzuordnen:

(α) Kongruente Mengen haben denselben Inhalt.

(β) Der Einheitswürfel hat den Inhalt 1.

(γ) Es ist $f(A+B) = f(A) + f(B)$.

(δ) Es ist $f(A+B+C+\ldots) = f(A) + f(B) + f(C) + \ldots$ für eine beschränkte Summe von abzählbar vielen Mengen.

Dieses Problem (unter Festhaltung des abzählbaren Summenaxioms (δ) und in Ausdehnung auf alle beschränkten Mengen) ist aber unlösbar, und zwar bereits für die linearen, um so mehr für die n-dimensionalen Mengen.[1] Wir führen den Beweis so, daß wir die Einheitsstrecke A $(0 \leqq x < 1)$, die den Inhalt $f(A) = 1$ haben soll, in abzählbar viele kongruente (genauer: zerlegungsgleiche) Summanden spalten, die also gleichen Inhalt haben müßten und damit das Axiom (δ) umstoßen. Lassen wir zunächst x alle reellen Zahlen durchlaufen und ordnen zwei Zahlen mit ganzzahliger Differenz denselben Punkt von A zu: geometrisch gesprochen, wir wickeln die gerade Linie, in der A liegt, auf eine Kreisperipherie vom Radius $\frac{1}{2\pi}$, die hierdurch umkehrbar eindeutig auf A bezogen wird. Zwei Mengen der Form

$$A_0 = \{x, y, z, \ldots\}, \quad A_1 = \{x+\alpha, y+\alpha, z+\alpha, \ldots\}$$

sind auf dem Kreise unmittelbar kongruent, da A_1 aus A_0 durch Verschiebung um den Bogen α entsteht. Auf der Einheitsstrecke sind sie zerlegungsgleich. Denn nimmt man $0 < \alpha < 1$ an, was offenbar erlaubt ist, zerlegt A_0 gemäß den Ungleichungen $0 \leqq x < 1-\alpha$ und $1-\alpha \leqq x < 1$ in $A_0 = B_0 + C_0$, hingegen A_1 gemäß den Ungleichungen $0 \leqq x < \alpha$ und $\alpha \leqq x < 1$ in $A_1 = C_1 + B_1$, so entsteht B_1 aus B_0 durch Verschiebung um α, C_1 aus C_0 durch Verschiebung um $\alpha - 1$; B_1 ist mit B_0, C_1 mit C_0 kongruent. Nach den Forderungen (α)(γ) haben also A_0 und A_1 denselben Inhalt.

Ist nun α eine irrationale Zahl, so entspricht jedem Punkt x eine abzählbare Menge

$$P_x = \{\ldots, x-2\alpha, x-\alpha, x, x+\alpha, x+2\alpha, \ldots\},$$

[1] Eine den obigen Ansprüchen genügende n-dimensionale Inhaltsbestimmung würde auch eine $(n-1)$-dimensionale liefern. Man lasse jeder $(n-1)$-dimensionalen Menge B den n-dimensionalen zylindrischen Körper A von der Basis B und der Höhe 1, d. h. die Menge derjenigen Punkte $x = (x_1, x_2, \ldots, x_{n-1}, x_n)$ entsprechen, für die $y = (x_1, x_2, \ldots, x_{n-1})$ ein Punkt von B und $0 \leqq x_n \leqq 1$ ist, und definiere $f_{n-1}(B) = f_n(A)$.

gewissermaßen die Menge der Ecken eines dem Kreise eingeschriebenen regulären Polygons, das aber unendlich viele Seiten hat und sich nicht schließt; zwei verschiedene Mengen P_x, P_y haben keinen Punkt gemein. Wir wählen jetzt aus jeder Menge P_x genau einen Punkt x aus und nennen $A_0 = \{x, y, \ldots\}$ die Menge dieser Punkte, ferner für ganzzahliges m

$$A_m = \{x + m\alpha,\ y + m\alpha,\ \ldots\}$$

die Menge, die aus A_0 auf dem Kreise durch Verschiebung um den Bogen $m\alpha$ entsteht. Es ist dann

$$A = \sum_m A_m = A_0 + A_1 + A_{-1} + A_2 + A_{-2} + \cdots$$

und alle A_m haben denselben Inhalt $f(A_m) = f(A_0)$. Das endliche Summenaxiom (γ) würde dann, weil für jede natürliche Zahl n

$$1 = f(A) \geqq f(A_1) + f(A_2) + \cdots + f(A_n) = n \cdot f(A_0)$$

ist, $f(A_0) = 0$ bedingen, und damit ist das abzählbare Summenaxiom (δ) verletzt.

Merkwürdigerweise ist selbst ohne die Forderung (δ) eine Lösung des Inhaltsproblems für alle beschränkten Mengen unmöglich, wenigstens im drei- oder mehrdimensionalen Raum (vgl. Anhang).

Die elementargeometrische Vorstufe, auf der die weiteren Inhaltsbegriffe beruhen, wollen wir nur flüchtig streifen. Als Elementarfigur verwenden wir auf der Geraden die Strecke, in der Ebene das Dreieck, im Raum das Tetraeder, im n-dimensionalen Raume das durch $n+1$ Punkte $a, b, \ldots, l$ bestimmte „Simplex", nämlich die Menge der Punkte x, deren rechtwinklige Koordinaten gegeben sind durch

$$x_i = \alpha a_i + \beta b_i + \cdots + \lambda l_i \quad (i = 1, 2, \ldots, n),$$

wobei die Zahlen $\alpha, \beta, \ldots, \lambda$ nichtnegativ und

$$\alpha + \beta + \cdots + \lambda = 1$$

ist. Dieser Menge[1] ordnen wir als Inhalt den durch $n!$ dividierten absoluten Betrag der Determinante

$$\begin{vmatrix} a_1 & b_1 & \cdots & l_1 \\ a_2 & b_2 & \cdots & l_2 \\ \cdot & \cdot & \cdots & \cdot \\ a_n & b_n & \cdots & l_n \\ 1 & 1 & \ldots & 1 \end{vmatrix}$$

zu; mit dieser Bewegungsinvariante als Grundlage alles folgenden ist die Lebesguesche Forderung (α) für kongruente Mengen wie

---

[1] Es ist die kleinste konvexe Menge (vgl. S. 329), welche die Punkte $a, b, \ldots, l$ enthält.

auch die Forderung (β) für den Einheitswürfel von vornherein garantiert.

Der bequemeren Redeweise wegen wählen wir die auf den Fall der Ebene bezüglichen Ausdrücke. Die genannten Mengen sind dann (abgeschlossene) Dreiecke D (inkl. Grenze), die sich auf Strecken oder Punkte reduzieren können; auch die Nullmenge wollen wir zu ihnen rechnen. Eine Summe

$$P = \mathfrak{S}(D_1, D_2, \ldots, D_n)$$

endlich vieler solcher Dreiecke heiße ein (abgeschlossenes) Polygon.[1] Da der Durchschnitt zweier Dreiecke ein Polygon ist, so ist Summe und Durchschnitt zweier Polygone wieder ein Polygon; die Polygone bilden einen Ring. Die obige Darstellung ist insbesondere so möglich, daß die Dreiecke paarweise keine inneren Punkte gemein haben; in diesem Falle definieren wir als Polygoninhalt die Summe der Dreiecksinhalte

$$f(P) = f(D_1) + f(D_2) + \cdots + f(D_n),$$

von welcher Zahl in bekannter Weise die Unabhängigkeit von der Dreieckszerlegung nachzuweisen ist. Sind P, Q abgeschlossene Polygone mit dem Durchschnitt R und der Summe S, so ist eine Darstellung

$$P = \mathfrak{S}(R, P'), \quad Q = \mathfrak{S}(R, Q'), \quad S = \mathfrak{S}(R, P', Q')$$

möglich, wo auch P', Q' abgeschlossene Polygone sind, die mit einander und mit R keinen inneren Punkt gemein haben; aus den Gleichungen $f(P) = f(R) + f(P')$ usw. folgt dann

$$f(P) + f(Q) = f(R) + f(S).$$

Wir wollen dies die Symmetrieformel der Polygoninhalte nennen, in Erinnerung daran, daß S, R die symmetrischen Grundmengen von P, Q sind (Kap. I, § 6).

§ 2. Der Peano-Jordansche Inhalt.

Im folgenden ist nur von beschränkten Mengen die Rede. Wir definieren:

Der äußere Inhalt $\varphi(A)$ einer Menge A ist die untere Schranke der Inhalte der einschließenden Polygone.

Wenn also P ein abgeschlossenes Polygon $\geqq A$ bedeutet, so

[1] Es kann in mehrere fremde Polygone zerfallen, in das Komplementärgebiet hineinragende Strecken enthalten und überhaupt noch sehr mannigfaltige Gestalten haben. Man bemerke, daß wir jetzt Polygonflächen, in Kap. VIII, § 11 Polygonumfänge als Polygone bezeichnen.

ist stets $f(P) \geqq \varphi(A)$, während sich zu jedem positiven ε mindestens ein P angeben läßt derart, daß

$$f(P) < \varphi(A) + \varepsilon;$$

wir wollen dies die ε-Bedingung für P nennen.

Für ein Polygon selbst ist $\varphi(P) = f(P)$.

Wenn jedes Polygon P, das A einschließt, zugleich A' einschließt, so ist $\varphi(A) \geqq \varphi(A')$. Insbesondere ist dies für $A \geqq A'$ der Fall. Aber es gilt auch für $A' = A_\alpha$, weil die P abgeschlossene Mengen sind; also ist $\varphi(A) \geqq \varphi(A_\alpha)$ und zugleich $\varphi(A_\alpha) \geqq \varphi(A)$, demnach

$$\varphi(A) = \varphi(A_\alpha).$$

Der äußere Inhalt einer Menge bleibt bei Hinzufügung ihrer Häufungspunkte ungeändert. Z. B. hat die Menge der rationalen Punkte des Einheitsquadrates den äußeren Inhalt 1.

Es seien A_1, A_2 zwei beschränkte Mengen,

$$A = \mathfrak{S}(A_1, A_2), \quad A' = \mathfrak{D}(A_1, A_2).$$

Wir schließen A_1, A_2 in Polygone P_1, P_2 gemäß der ε-Bedingung ein und setzen

$$P = \mathfrak{S}(P_1, P_2), \quad P' = \mathfrak{D}(P_1, P_2).$$

Da $P \geqq A$, $P' \geqq A'$, so erhält man nach der Symmetrieformel der Polygoninhalte

$$\begin{aligned} \varphi(A_1) + \varphi(A_2) + 2\varepsilon &> f(P_1) + f(P_2) \\ &= f(P) + f(P') \\ &\geqq \varphi(A) + \varphi(A'), \end{aligned}$$

also

$$\varphi(A) + \varphi(A') < \varphi(A_1) + \varphi(A_2) + 2\varepsilon$$

für jedes positive ε und demnach

(1) $$\varphi(A) + \varphi(A') \leqq \varphi(A_1) + \varphi(A_2),$$

was man wieder als Symmetrieformel der äußeren Inhalte bezeichnen kann.

Wir heben einen Fall hervor, wo das Gleichheitszeichen gilt. Wenn sich A_1, A_2 in Polygone Q_1, Q_2 einschließen lassen, die keinen innern Punkt gemein haben, so schließe man A in ein Polygon P gemäß der ε-Bedingung ein; die Polygone $P_1 = \mathfrak{D}(P, Q_1)$, $P_2 = \mathfrak{D}(P, Q_2)$ schließen A_1, A_2 ein und haben keinen inneren Punkt gemein. Wegen $P \geqq \mathfrak{S}(P_1, P_2)$ ist

$$\varphi(A) + \varepsilon > f(P) \geqq f(P_1) + f(P_2) \geqq \varphi(A_1) + \varphi(A_2),$$

also $\varphi(A) \geqq \varphi(A_1) + \varphi(A_2)$ und in Verbindung mit (1)

$$\varphi(A) = \varphi(A_1) + \varphi(A_2).$$

Insbesondere gilt dies, wenn A_1 in einem Polygon P liegt, mit dem

A_2 keinen Punkt gemein hat ($A_1 \leqq P$, $A_2 \leqq E - P$). Sind daher $A \leqq B$ zwei Mengen, zwischen denen ein Polygon liegt ($A \leqq P \leqq B$), so ist

$$B - A = (B - P) + (P - A),$$
$$\varphi(B - A) = \varphi(B - P) + \varphi(P - A),$$

insbesondere (für $A = 0$)

$$\varphi(B) = \varphi(B - P) + \varphi(P),$$

also

$$\varphi(B) - \varphi(B - A) = \varphi(P) - \varphi(P - A).$$

Diese Formel zeigt, daß der Ausdruck

$$\psi(A) = \varphi(P) - \varphi(P - A),$$

wo P ein Polygon über A ist, von P unabhängig ist; denn sind P, P' zwei solche Polygone und $B \geqq \mathfrak{S}(P, P')$, so ist

$$\varphi(P) - \varphi(P - A) = \varphi(P') - \varphi(P' - A) = \varphi(B) - \varphi(B - A).$$

Diese nur von A abhängige Zahl $\psi(A)$ heißt der **innere Inhalt** von A; da nach (1) $\varphi(P) \leqq \varphi(A) + \varphi(P - A)$, so ist $\psi(A) \leqq \varphi(A)$, der innere Inhalt höchstens dem äußeren gleich.

Für ein Polygon P ist, wenn wir es in sich selbst einschließen, $\psi(P) = \varphi(P) - \varphi(0) = \varphi(P) = f(P)$. Allgemein definieren wir, wenn äußerer und innerer Inhalt übereinstimmen,

$$f(A) = \varphi(A) = \psi(A)$$

als **Inhalt** schlechthin und nennen die Menge A **meßbar** (im **Peano-Jordanschen** Sinn).

Sind wieder A_1, A_2 zwei beliebige Mengen, A ihre Summe, A' ihr Durchschnitt, ferner B_1, B_2, B, B' die Komplemente dieser Mengen in bezug auf ein A einschließendes Polygon P, wobei also B der Durchschnitt und B' die Summe von B_1, B_2 ist, so gibt die Symmetrieformel der äußeren Inhalte

$$\varphi(B) + \varphi(B') \leqq \varphi(B_1) + \varphi(B_2)$$

durch Subtraktion von $2\varphi(P)$ die **Symmetrieformel der inneren Inhalte**

(2) $$\psi(A) + \psi(A') \geqq \psi(A_1) + \psi(A_2).$$

Für zwei fremde Mengen A_1, A_2 gelten die **Summenformeln**

(3) $$\left\{\begin{array}{l} \varphi(A_1 + A_2) \leqq \varphi(A_1) + \varphi(A_2) \\ \varphi(A_1 + A_2) \geqq \psi(A_1) + \varphi(A_2) \\ \psi(A_1 + A_2) \leqq \psi(A_1) + \varphi(A_2) \\ \psi(A_1 + A_2) \geqq \psi(A_1) + \psi(A_2) \end{array}\right.$$

von denen die erste und vierte eine Spezialisierung der Symmetrieformeln sind. Die mittleren erhält man folgendermaßen. Wir

schließen $A_1 + A_2$ in ein Polygon P ein und es sei $P = A_1 + A_2 + A_3$. Auf Grund der Beziehung zwischen äußerem und innerem Inhalt und mit Benutzung der beiden schon bewiesenen Formeln (3) ist dann

$$\varphi(A_1 + A_2) - \varphi(A_2) = \psi(A_1 + A_3) - \psi(A_3) \geqq \psi(A_1),$$
$$\psi(A_1 + A_2) - \psi(A_1) = \varphi(A_2 + A_3) - \varphi(A_3) \leqq \varphi(A_2).$$

Aus den mittleren Summenformeln erhält man noch zwei weitere durch Vertauschung der Indices 1, 2.

Sind jetzt A_1, A_2 zwei meßbare Mengen mit der Summe A und dem Durchschnitt A', so folgt aus den Symmetrieformeln

$$\psi(A) + \psi(A') \geqq f(A_1) + f(A_2) \geqq \varphi(A) + \varphi(A'),$$

also, weil stets $\psi \leqq \varphi$, die Meßbarkeit von A und A' nebst der Symmetrieformel der Inhalte

$$f(A) + f(A') = f(A_1) + f(A_2).$$

Summe und Durchschnitt von zwei (oder endlich vielen) meßbaren Mengen sind meßbar.

Ist ferner $A_1 + A_2$ meßbar, so folgt aus den mittleren Summenformeln

$$f(A_1 + A_2) = \psi(A_1) + \varphi(A_2) = \varphi(A_1) + \psi(A_2),$$

was zunächst eine Verallgemeinerung der Formel ergibt, durch die der innere Inhalt definiert wurde: es ist

$$\psi(A) = f(M) - \varphi(M - A),$$

wenn M eine meßbare Menge $\geqq A$ bedeutet. Außerdem folgt, wenn A_1 meßbar ist, auch die Meßbarkeit von A_2: eine Differenz meßbarer Mengen ist meßbar. Die meßbaren Mengen bilden einen Körper.

Ist A_1 meßbar, so geben die Summenformeln

$$\varphi(A_1 + A_2) = f(A_1) + \varphi(A_2), \quad \psi(A_1 + A_2) = f(A_1) + \psi(A_2).$$

Insbesondere haben zwei Mengen, deren Differenz den (äußeren) Inhalt Null hat, denselben äußeren und denselben inneren Inhalt.

Ein abgeschlossenes Dreieck D hat denselben Inhalt wie sein Inneres $\Delta = D_i$, da der Rand D_r vom Inhalt Null (ein zur Streckensumme ausgeartetes Polygon) ist. Wir nennen die Δ (wozu auch die Nullmenge gehört) Dreiecksgebiete und ihre Summen

$$\Pi = \mathfrak{S}(\Delta_1, \Delta_2, \ldots, \Delta_n)$$

Polygongebiete; da der Durchschnitt zweier Δ ein Π ist, so bilden auch die Polygongebiete einen Ring. Sie sind meßbar und ihre Inhalte genügen natürlich wie die aller meßbaren Mengen der Symmetrieformel.

Durch Tilgung eines Δ aus einem D entsteht, wie man leicht sieht, ein abgeschlossenes Polygon; $D - \Delta$ ist ein P, ebenso $\Delta - D$ ein Π. $D - \Pi$ ist Durchschnitt von Mengen $D - \Delta$, also ein P, und $\Delta - P$ ein Π. $P - \Pi$ oder allgemeiner $P - \mathfrak{D}(P, \Pi)$ ist ein P, als Durchschnitt $\mathfrak{D}(P, D - \Pi)$ für irgend ein $D \geqq \mathfrak{S}(P, \Pi)$; $\Pi - P$ oder $\Pi - \mathfrak{D}(P, \Pi)$ ist ein Π.

Hiernach ist leicht zu erkennen, daß der innere Inhalt $\psi(A)$ einer Menge die obere Schranke der Inhalte aller eingeschlossenen Polygongebiete ist, welche Eigenschaft auch zur Definition hätte dienen können. Denn schließen wir A in ein Dreiecksgebiet Δ ein und setzen $\Delta = A + B$, so ist $\psi(A) = f(\Delta) - \varphi(B)$. Schließen wir B in ein P mit $f(P) < \varphi(B) + \varepsilon$ ein, so ist

$$\Pi = \Delta - \mathfrak{D}(\Delta, P) \leqq A$$

ein Polygongebiet und

$$f(\Pi) \geqq f(\Delta) - f(P) > \psi(A) - \varepsilon,$$

während andererseits, für jedes $\Pi \leqq A$, $\psi(A) \geqq \psi(\Pi) = f(\Pi)$ ist; also ist $\psi(A)$ die obere Schranke der $f(\Pi)$. Da jedes Polygongebiet $\leqq A$ auch im Innern A_i von A liegt, so folgt als Analogon zur Formel $\varphi(A) = \varphi(A_\alpha)$ die Formel

$$\psi(A) = \psi(A_i).$$

Der innere Inhalt bleibt bei Weglassung der Randpunkte ungeändert; eine Menge ohne innere Punkte (Randmenge) hat stets den inneren Inhalt Null. Z. B. hat die Menge der rationalen Punkte des Einheitsquadrats den inneren Inhalt 0; da sie den äußeren Inhalt 1 hatte, ist sie nicht meßbar. Das gleiche gilt von der Menge der irrationalen Punkte. Da jedes Gebiet insichdicht, also A_i im Kern A_k von A enthalten ist, so ist auch $\psi(A_k) = \psi(A)$; die Kohärenzen A_h, A_{hh}, ... und im Falle einer abgeschlossenen Menge die Ableitungen haben alle denselben inneren Inhalt wie A selbst.[1] Da ferner jeder innere Punkt von A auch Verdichtungspunkt, also A_i in der Menge $A_v = \mathfrak{D}(A, A_\gamma)$ enthalten ist, so ist auch $\psi(A_v) = \psi(A)$; eine höchstens abzählbare Menge hat stets den inneren Inhalt Null.

Ein Polygongebiet Π hat denselben Inhalt wie das abgeschlossene Polygon Π_α, das aus Π durch Hinzufügung der Häufungspunkte entsteht, ein abgeschlossenes Polygon P denselben Inhalt wie sein Inneres, das Polygongebiet P_i.[2] Andererseits gibt es zu

---

[1] Für den Fall einer abgeschlossenen Menge gilt dies auch vom äußeren Inhalt (vgl. S. 414).

[2] Für $\Pi = \mathfrak{S}\,\Delta_m$ ist $\Pi_\alpha = \mathfrak{S}\,\Delta_{m\alpha}$ ein abgeschlossenes Polygon. Daß P_i ein Polygongebiet ist, kann man so unmittelbar nicht schließen, da aus $P = \mathfrak{S} D_m$ nur $P_i \geqq \mathfrak{S}\, D_{mi}$ folgt. Schließt man aber P in ein Δ ein und setzt $\Delta - P = \Pi$, so ist $P_i = \Delta - \mathfrak{D}(\Delta, \Pi_\alpha)$ ein Polygongebiet.

jedem D ein einschließendes $\varDelta$ mit beliebig kleinem $f(\varDelta - D)$, also zu jedem $P = \mathfrak{S} D_m$ ein einschließendes

$$\varPi = \mathfrak{S} \varDelta_m, \quad \varPi - P \leqq \mathfrak{S}(\varDelta_m - D_m)$$

mit beliebig kleinem $f(\varPi - P)$, und ebenso zu jedem $\varPi$ ein eingeschlossenes P mit beliebig kleinem $f(\varPi - P)$. Man kann also ein P durch ein eingeschlossenes $\varPi$ genau, durch ein einschließendes $\varPi$ beliebig nahe approximieren, und analog ein $\varPi$ durch ein einschließendes resp. eingeschlossenes P. Danach können die P und $\varPi$ einander vertreten; $\varphi(A)$ ist auch die untere Schranke der Inhalte der einschließenden $\varPi$, $\psi(A)$ die obere Schranke der Inhalte der eingeschlossenen P.

§ 3. Das Lebesguesche Maß.

Wir geben, indem wir nach wie vor mit beschränkten Mengen operieren, eine zweite Definition der Zahlen $\varphi(A)$, $\psi(A)$, die wir jetzt äußeres und inneres Maß und im Falle der Gleichheit das Maß $f(A)$ der Menge A nennen. Die im vorigen Paragraphen erklärten Inhalte sollen von nun an mit $\varphi_0(A)$, $\psi_0(A)$, $f_0(A)$ bezeichnet werden, wobei der Index 0 an $f_0(A)$ indessen auch wegbleiben kann. Die Mengen, für die $f(A)$ existiert, heißen meßbar (im Lebesgueschen Sinne).

An Stelle der Polygone treten, für die Bestimmung des äußeren Maßes, jetzt die Summen von Polygonfolgen, und zwar wollen wir hier Polygongebiete verwenden. Unter

$$G = \mathfrak{S}(\varPi_1, \varPi_2, \ldots) = \mathfrak{S} \varPi_n$$

verstehen wir die beschränkte Summe einer Folge aufsteigender Polygongebiete $\varPi_1 \leqq \varPi_2 \leqq \ldots$; eine solche Polygonsumme ist ein Gebiet, wie umgekehrt jedes Gebiet sich als Summe der in ihm enthaltenen Dreiecksgebiete mit rationalen Eckpunkten, also als Summe einer Dreiecksfolge $G = \mathfrak{S}(\varDelta_1, \varDelta_2, \ldots)$ und demnach, mit $\varPi_n = \mathfrak{S}(\varDelta_1, \varDelta_2, \ldots, \varDelta_n)$, als Summe einer aufsteigenden Folge von Polygongebieten darstellen läßt. Wir ordnen jener Polygonsumme die Zahl

$$g = \lim \pi_n = \lim f(\varPi_n)$$

zu, von der wir natürlich zunächst noch gar nicht wissen, ob sie bloß von der Menge G und nicht auch von der gewählten Summendarstellung abhängt. Wir definieren:

Das äußere Maß $\varphi(A)$ ist die untere Schranke der g für alle die Menge A einschließenden Gebiete G.

Es ist also für $G \geqq A$ stets $g \geqq \varphi(A)$, während sich zu jedem positiven ε ein G mit

$$g < \varphi(A) + \varepsilon$$

angeben läßt.

Für $A \geqq A'$ ist offenbar $\varphi(A) \geqq \varphi(A')$.

Es ist stets

$$\varphi(A) \leqq \varphi_0(A),$$

das äußere Maß höchstens gleich dem äußeren Inhalt. Denn da $\varphi_0(A)$ auch die untere Schranke aller $f(\Pi)$ für $\Pi \geqq A$ ist (§ 2 am Ende), so können wir ein $\Pi \geqq A$ mit $f(\Pi) < \varphi_0(A) + \varepsilon$ bestimmen; die Polygonsumme $\mathfrak{S}(\Pi, \Pi, \Pi, \ldots)$ gibt dann $f(\Pi) \geqq \varphi(A)$, also $\varphi_0(A) + \varepsilon > \varphi(A)$ und damit die behauptete Ungleichung.

Für eine abgeschlossene Menge A ist $\varphi(A) = \varphi_0(A)$.

Denn bestimmen wir $G \geqq A$ nach der ε-Bedingung. Die Menge A ist in $\mathfrak{S}(\Pi_1, \Pi_2, \ldots)$, also nach dem Borelschen Satze schon in einer endlichen Summe $\mathfrak{S}(\Pi_1, \Pi_2, \ldots, \Pi_n) = \Pi_n$ enthalten, demnach

$$\varphi_0(A) \leqq \pi_n \leqq \lim \pi_n < \varphi(A) + \varepsilon,$$

also $\varphi_0(A) \leqq \varphi(A)$ und in Verbindung mit der allgemeinen Formel das Gleichheitszeichen.

Insbesondere ist, für ein abgeschlossenes Polygon,

$$\varphi(P) = \varphi_0(P) = f(P)$$

gleich dem elementaren Inhalt.

Sind A_1, A_2 zwei Mengen mit der Summe A und dem Durchschnitt A', so schließen wir sie in zwei Gebiete

$$G_1 = \mathfrak{S}\, \Pi_{1n}, \quad G_2 = \mathfrak{S}\, \Pi_{2n}$$

gemäß der ε-Bedingung ein und setzen

$$G = \mathfrak{S}(G_1, G_2), \qquad G' = \mathfrak{D}(G_1, G_2),$$
$$\Pi_n = \mathfrak{S}(\Pi_{1n}, \Pi_{2n}), \quad \Pi_n' = \mathfrak{D}(\Pi_{1n}, \Pi_{2n}).$$

Dann ist

$$G = \mathfrak{S}\, \Pi_n, \quad G' = \mathfrak{S}\, \Pi_n';$$

die erste Formel ist evident, die zweite entsteht aus

$$G' = \mathfrak{S}\, \mathfrak{D}(\Pi_{1m}, \Pi_{2n})$$

durch Weglassung derjenigen Summanden, die wegen des Aufsteigens der Folgen Π_{1n}, Π_{2n} in den Summanden Π_n' enthalten sind.

Da nun die Π_n, Π_n' wieder aufsteigende Folgen bilden, so ist nach der Symmetrieformel für die Inhalte der Polygongebiete

$$\begin{aligned} \varphi(A_1) + \varphi(A_2) + 2\varepsilon &> \lim(\pi_{1n} + \pi_{2n}) \\ &= \lim(\pi_n + \pi_n') \\ &\geqq \varphi(A) + \varphi(A'), \end{aligned}$$

also gilt auch hier die Symmetrieformel der äußeren Maße

(1) $$\varphi(A) + \varphi(A') \leqq \varphi(A_1) + \varphi(A_2).$$

Wir zeigen wieder, daß in einem gewissen Fall das Gleichheitszeichen gilt. Wenn sich A_1, A_2 in Polygongebiete Γ_1, Γ_2 ohne gemeinsamen Punkt einschließen lassen, so schließe man A in ein Gebiet $G = \mathfrak{S}\,\Pi_n$ gemäß der ε-Bedingung ein. Setzt man

$$\Pi_{1n} = \mathfrak{D}(\Pi_n, \Gamma_1), \quad \Pi_{2n} = \mathfrak{D}(\Pi_n, \Gamma_2),$$

so bilden diese Polygongebiete aufsteigende Folgen, deren Summen A_1 und A_2 einschließen; also ist

$$\varphi(A) + \varepsilon > \lim \pi_n \geqq \lim (\pi_{1n} + \pi_{2n}) \geqq \varphi(A_1) + \varphi(A_2)$$

und demgemäß

$$\varphi(A) = \varphi(A_1) + \varphi(A_2).$$

Wenn A_1 in einem abgeschlossenen Polygon P liegt, mit dem A_2 keinen Punkt gemein hat ($A_1 \leqq P$, $A_2 \leqq E - P$), so gilt diese Formel auch noch. Denn die Menge $B = \mathfrak{D}(A_1, P_r)$ der Punkte von A_1, die auf dem Rande von P liegen, hat den äußeren Inhalt 0 und erst recht das äußere Maß 0; nach (1) ist $\varphi(A) \leqq \varphi(A-B) + \varphi(B) = \varphi(A-B)$, also $\varphi(A-B) = \varphi(A)$ und ebenso $\varphi(A_1 - B) = \varphi(A_1)$. Die Menge $A_1 - B = \mathfrak{D}(A_1, P_i)$ liegt in dem Polygongebiet P_i, A_2 in dem dazu fremden Polygongebiet $\Delta - P$, wenn Δ ein Dreiecksgebiet $\geqq P + A_2$ ist, also ist

$$\varphi(A - B) = \varphi(A_1 - B) + \varphi(A_2),$$

$$\varphi(A) = \varphi(A_1) + \varphi(A_2).$$

Hieraus folgt nun genau wie in § 2, daß der Ausdruck

$$\psi(A) = \varphi(P) - \varphi(P - A),$$

wenn P ein abgeschlossenes Polygon über A ist, von P unabhängig ist; wir nennen ihn das innere Maß von A. Die Vergleichung mit dem inneren Inhalt

$$\psi_0(A) = \varphi_0(P) - \varphi_0(P - A),$$

wobei $\varphi(P) = \varphi_0(P)$ und $\varphi(P - A) \leqq \varphi_0(P - A)$, gibt $\psi(A) \geqq \psi_0(A)$; das innere Maß ist mindestens gleich dem inneren Inhalt. Nach (1) ist ferner $\psi(A) \leqq \varphi(A)$, sodaß zwischen den vier definierten Zahlen die Beziehung

$$\psi_0(A) \leqq \psi(A) \leqq \varphi(A) \leqq \varphi_0(A)$$

besteht. Danach existiert das Maß jedenfalls, wenn der Inhalt existiert, und alle im Peano-Jordanschen Sinne meßbaren Mengen sind es auch im Lebesgueschen Sinne; wogegen das Umgekehrte, wie wir sehen werden, nicht der Fall ist, der Kreis der meßbaren Mengen also durch die Lebesguesche Maßbestimmung wirklich erweitert wird.

Wörtlich wie in § 2 beweist man auch die Symmetrieformel der inneren Maße

(2) $$\psi(A) + \psi(A') \geqq \psi(A_1) + \psi(A_2)$$

und die Summenformeln

(3) $$\begin{cases} \varphi(A_1 + A_2) \leqq \varphi(A_1) + \varphi(A_2) \\ \varphi(A_1 + A_2) \geqq \psi(A_1) + \varphi(A_2) \\ \psi(A_1 + A_2) \leqq \psi(A_1) + \varphi(A_2) \\ \psi(A_1 + A_2) \geqq \psi(A_1) + \psi(A_2) \end{cases}$$

da dies alles ja, unabhängig von der Bedeutung von $\varphi(A)$, nur auf der Symmetrieformel (1) und der Definitionsformel von $\psi(A)$ beruht. Summe, Durchschnitt und Differenz zweier meßbarer Mengen ist wieder meßbar; die meßbaren Mengen bilden einen Körper.

Aber jetzt kommt etwas Neues hinzu: die meßbaren Mengen bilden ein σ-System (wenigstens in dem Sinne, daß die beschränkte Summe einer Folge meßbarer Mengen wieder meßbar ist) und ein δ-System. Es gilt nämlich:

Ist $A = \mathfrak{S} A_m$ die Summe einer aufsteigenden Folge $A_1 \leqq A_2 \leqq \ldots$, so ist

(4) $$\varphi(A) = \lim \varphi(A_m).$$

Zunächst ist jedenfalls $\varphi(A) \geqq \varphi(A_m)$, also $\varphi(A) \geqq \lim \varphi(A_m)$. Schließen wir nun A_m in ein Gebiet $G_m = \mathfrak{S}(\Pi_{m1}, \Pi_{m2}, \ldots)$ mit

$$g_m = \lim_n \pi_{mn} < \varphi(A_m) + \varepsilon$$

ein. Wir behaupten, daß die Mengen Π_{mn} nicht nur mit dem Index n, sondern auch mit dem Index m aufsteigend ($\Pi_{1n} \leqq \Pi_{2n} \leqq \ldots$) gewählt werden können. Nachdem G_1 gewählt ist, schließe man A_2 in G_2 mit der Ungleichung

$$g_2 < \varphi(A_2) + [\varphi(A_1) + \varepsilon - g_1]$$

ein und setze wieder

$$G = \mathfrak{S}(G_1, G_2), \quad G' = \mathfrak{D}(G_1, G_2)$$

usw. wie beim Beweise der Symmetrieformel (1). Dabei ist

$$G \geqq \mathfrak{S}(A_1, A_2) = A_2, \quad G' \geqq \mathfrak{D}(A_1, A_2) = A_1,$$

also

$$\begin{aligned} g + \varphi(A_1) \leqq g + g' &= g_1 + g_2 \\ &< \varphi(A_1) + \varphi(A_2) + \varepsilon, \\ g &< \varphi(A_2) + \varepsilon. \end{aligned}$$

Das Gebiet $G = \mathfrak{S}\,\Pi_n$ mit $\Pi_n = \mathfrak{S}(\Pi_{1n}, \Pi_{2n})$ erfüllt also die Forderungen, die wir an G_2 stellen wollten: es schließt A_2 ein mit $g < \varphi(A_2) + \varepsilon$ und es ist $\Pi_{1n} \leqq \Pi_n$. Ändern wir die Bezeichnung

und nennen G jetzt G_2, so ist $g_2 < \varphi(A_2) + \varepsilon$ und $\Pi_{1n} \leqq \Pi_{2n}$. Verfahren wir mit G_3 gegenüber G_2, wie wir mit G_2 gegenüber G_1 verfahren sind, und setzen dies fort, so ist die obige Behauptung bewiesen.

Die Summe aller G_m läßt sich nunmehr unter Fortlassung entbehrlicher Summanden in der Gestalt

$$G = \mathfrak{S}\, G_m = \mathfrak{S}\, \Pi_{mn} = \mathfrak{S}\, \Pi_{mm}$$

schreiben, wo die Π_{mm} wieder eine aufsteigende Folge bilden. Man hat

$$\pi_{mm} \leqq \lim_n \pi_{mn} < \varphi(A_m) + \varepsilon$$

und wegen $A \leqq G$

$$\varphi(A) \leqq \lim \pi_{mm} \leqq \lim \varphi(A_m) + \varepsilon,$$

also $\varphi(A) \leqq \lim \varphi(A_m)$ und in Verbindung mit der umgekehrten Ungleichung die behauptete Gleichung (4). Der Beweis ist, wie man sieht, wieder einmal ein „Diagonalverfahren“, das eine Doppelfolge in eine einfache Folge umwandelt.

Für die inneren Maße unserer Mengen gilt $\psi(A) \geqq \psi(A_m)$ und

$$\psi(A) \geqq \lim \psi(A_m);$$

hier ist das Ungleichheitszeichen im allgemeinen nicht entbehrlich, wie wir aus einem Beispiel (S. 419) sehen werden. Sind nun die Mengen A_m meßbar, so folgt

$$\varphi(A) = \lim f(A_m) \leqq \psi(A)$$

und daraus die Meßbarkeit von A nebst $f(A) = \lim f(A_m)$.

Durch Komplementbildung in bezug auf ein Polygon oder eine meßbare Menge gewinnt man den dualistischen Satz:

Ist $A = \mathfrak{D}\, A_m$ der Durchschnitt einer absteigenden Folge $A_1 \geqq A_2 \geqq \ldots$, so ist

(5) $$\psi(A) = \lim \psi(A_m).$$

Gleichzeitig ist

$$\varphi(A) \leqq \lim \varphi(A_m),$$

woraus man wieder im Falle der Meßbarkeit der Mengen A_m die Meßbarkeit von A nebst $f(A) = \lim f(A_m)$ folgert. Also:

Die (beschränkte) Summe und der Durchschnitt einer Folge meßbarer Mengen ist meßbar.

Dieser Satz bezeichnet den großen Fortschritt der Lebesgueschen Maßbestimmung gegenüber der Peano-Jordanschen Inhaltsbestimmung.

Jedes Gebiet ist meßbar, denn $G = \mathfrak{S}\, \Pi_n$ ist Summe einer Folge meßbarer Mengen (Polygongebiete). Zugleich ist, bei aufsteigenden Π_n, $f(G) = \lim f(\Pi_n)$; die zuvor benutzten Zahlen g stellen sich also

als die Maße der betreffenden Gebiete heraus, und das äußere Maß $\varphi(A)$ ist die untere Schranke der Maße der A einschließenden Gebiete.

Jede abgeschlossene Menge F ist meßbar, denn in bezug auf ein einschließendes Dreiecksgebiet ist das Komplement $\Delta - F = G$ ein Gebiet, F eine Differenz von Gebieten. Man erkennt auch auf gleiche Weise, daß die abgeschlossenen Mengen mit den Durchschnitten $\mathfrak{D} P_n$ absteigender Folgen abgeschlossener Polygone identisch sind, daß, für $F = \mathfrak{D} P_n$, $f = f(F) = \lim p_n = \lim f(P_n)$ ist, und daß das innere Maß $\psi(A)$ die obere Schranke der Zahlen f für $F \leqq A$ oder die obere Schranke der Maße der in A enthaltenen abgeschlossenen Mengen ist.

Für eine abgeschlossene Menge hatten wir $f(F) = \varphi(F) = \varphi_0(F)$ gefunden, das Maß gleich dem äußeren Inhalt; schließen wir ein Gebiet G in ein D ein, so daß $D - G = F$ abgeschlossen ist, so ist $f(G) = f(D) - \varphi_0(F) = \psi_0(G)$, das Maß eines Gebietes gleich dem inneren Inhalt. Wir erhalten also folgende übersichtliche Zusammenstellung (Young):

äußerer Inhalt = untere Schranke der Inhalte der einschließenden Polygone,
innerer Inhalt = obere Schranke der Inhalte der eingeschlossenen Polygone,
äußeres Maß = untere Schranke der inneren Inhalte der einschließenden Gebiete,
inneres Maß = obere Schranke der äußeren Inhalte der eingeschlossenen abgeschlossenen Mengen.

Zu den meßbaren Mengen gehören ferner alle die, die aus den F und G durch Summen- und Durchschnittsbildung von Folgen entstehen, also die Borelschen (oder, wie man auch sagt, im Borelschen Sinne meßbaren) Mengen F_σ, G_δ, $F_{\sigma\delta}$, $G_{\delta\sigma}$, ...

Eine abzählbare Menge hat, als Summe einer aufsteigenden Folge endlicher Mengen, das Maß Null. Die Menge der rationalen Punkte im Einheitsquadrat hat das Maß 0, die der irrationalen das Maß 1; beide Mengen sind im Lebesgueschen, nicht aber im Peano-Jordanschen Sinne meßbar.

Jede Menge A kann von oben her genau durch eine Menge $G_\delta \geqq A$, von unten her genau durch eine Menge $F_\sigma \leqq A$ approximiert werden, d. h. so daß

$$\varphi(A) = f(G_\delta), \quad \psi(A) = f(F_\sigma).$$

Denn wählt man ein $G_n \geqq A$ mit $f(G_n) < \varphi(A) + \frac{1}{n}$, so hat der Durchschnitt G_δ dieser Gebietsfolge das Maß $\varphi(A)$; analog ist F_σ zu bilden.

Eine Menge mit positivem inneren Maß ist von der Mächtigkeit des Kontinuums, da das approximierende F_σ weder endlich noch abzählbar sein kann (S. 321).

Da $A_\alpha - A_i = A_g$ die Grenze von A und abgeschlossen ist, so hat man

$$\varphi_0(A) = \varphi_0(A_\alpha) = f(A_\alpha),$$
$$\psi_0(A) = \psi_0(A_i) = f(A_i),$$
$$\varphi_0(A) - \psi_0(A) = \quad f(A_g) = \varphi_0(A_g);$$

der Unterschied zwischen dem äußeren und inneren Inhalt von A ist das Maß oder der äußere Inhalt der Grenze von A. Eine Menge hat Inhalt, wenn ihre Grenze das Maß oder den Inhalt Null hat.

Ist F abgeschlossen und $A \leqq F$ höchstens abzählbar, so ist

$$\varphi_0(F) = f(F) = f(F - A) \leqq \varphi_0(F - A) \leqq \varphi_0(F),$$

also $\varphi_0(F) = \varphi_0(F - A)$. Der äußere Inhalt einer abgeschlossenen Menge wird durch Weglassung von endlich oder abzählbar vielen Punkten nicht geändert, ist also (wie der innere Inhalt) dem äußeren Inhalt aller Ableitungen und des Kernes gleich.

Für die (beschränkte) Summe $S = \mathfrak{S}(A_1, A_2, \ldots)$ einer beliebigen Mengenfolge gilt, weil S die Summe der aufsteigenden Folge von Mengen $\mathfrak{S}(A_1, A_2, \ldots, A_n)$ ist,

$$\varphi(S) = \lim \varphi\left(\mathfrak{S}(A_1, A_2, \ldots, A_n)\right)$$
$$\leqq \lim [\varphi(A_1) + \varphi(A_2) + \cdots + \varphi(A_n)],$$

(6) $$\varphi(S) \leqq \varphi(A_1) + \varphi(A_2) + \cdots,$$

welche Formel natürlich nur bei Konvergenz der rechten Seite einen Sinn hat. Für paarweise fremde Mengen A_n gilt außerdem

$$\psi(S) \geqq \lim \psi(A_1 + A_2 + \cdots + A_n)$$
$$\geqq \lim [\psi(A_1) + \psi(A_2) + \cdots + \psi(A_n)],$$

(7) $$\psi(S) \geqq \psi(A_1) + \psi(A_2) + \cdots.$$

Also im Falle meßbarer Mengen A_n

$$f(A_1 + A_2 + \cdots) = f(A_1) + f(A_2) + \cdots.$$

Ferner hat man im allgemeinen Falle für den oberen und unteren Limes der Mengenfolge (S. 21)

$$M = \operatorname{Lim\,sup} A_n = \mathfrak{D}(S_1, S_2, \ldots), \qquad L = \operatorname{Lim\,inf} A_n = \mathfrak{S}(D_1, D_2, \ldots)$$

mit

$$S_n = \mathfrak{S}(A_n, A_{n+1}, \ldots), \; D_n = \mathfrak{D}(A_n, A_{n+1}, \ldots),$$

also, da die S_n eine absteigende, die D_n eine aufsteigende Folge bilden,

$$\psi(M) = \lim \psi(S_n), \quad \varphi(L) = \lim \varphi(D_n).$$

Ist z. B. $M = 0$ oder vom innern Maße Null, so folgt $\lim \psi(S_n) = 0$ und erst recht $\lim \psi(A_n) = 0$.

Wir fragen noch, wann die Ungleichungen für das äußere und innere Maß einer Summe in Gleichungen übergehen. Nennen wir den Durchschnitt von A mit einer meßbaren Menge eine in A meßbare Menge; dieser Begriff der relativen Meßbarkeit ist ein Analogon der relativen Abgeschlossenheit und anderer Relativbegriffe. Ist nun M eine meßbare Menge, $D = \mathfrak{D}(A, M)$ in A meßbar und $S = \mathfrak{S}(A, M) = M + (A - D)$, so ist nach Symmetrie- und Summenformeln

$$\varphi(A) + \varphi(M) \geqq \varphi(D) + \varphi(S)$$
$$\geqq \varphi(D) + \varphi(A - D) + \psi(M),$$

also, wegen der Meßbarkeit von M, $\varphi(A) \geqq \varphi(D) + \varphi(A - D)$ und schließlich

$$\varphi(A) = \varphi(D) + \varphi(A - D),$$

ebenso

$$\psi(A) = \psi(D) + \psi(A - D).$$

Daraus folgt:

Wird $A = A_1 + A_2 + \ldots$ in endlich oder abzählbar viele, paarweise fremde, in A meßbare Mengen A_n zerlegt, so ist

$$\varphi(A) = \varphi(A_1) + \varphi(A_2) + \cdots,$$
$$\psi(A) = \psi(A_1) + \psi(A_2) + \cdots.$$

Ist nämlich $A_n = \mathfrak{D}(A, M_n)$ in A meßbar, so hat man zunächst

$$\varphi(A) = \varphi(A_1) + \varphi(A - A_1),$$

sodann, weil $A_2 = \mathfrak{D}(A - A_1, A_2) = \mathfrak{D}(A - A_1, M_2)$ auch in $A - A_1$ meßbar ist,

$$\varphi(A - A_1) = \varphi(A_2) + \varphi(A - A_1 - A_2),$$

und so fortfahrend für $B_n = A_1 + A_2 + \ldots + A_n$:

$$\varphi(A) = \varphi(A_1) + \varphi(A_2) + \ldots + \varphi(A_n) + \varphi(A - B_n),$$
$$\psi(A) = \psi(A_1) + \psi(A_2) + \ldots + \psi(A_n) + \psi(A - B_n).$$

Daraus folgt die Behauptung für endlich viele Summanden; für abzählbar viele schließt man einerseits

$$\varphi(A) \geqq \varphi(A_1) + \varphi(A_2) + \cdots$$

und in Verbindung mit (6) das Gleichheitszeichen, andererseits

$$\psi(A) = \psi(A_1) + \psi(A_2) + \cdots,$$

da die absteigende Folge der $A - B_n$ den Durchschnitt 0 hat und nach (5) $\lim \psi(A - B_n) = 0$ ist.

Ein Teil dieses Satzes ist umkehrbar: wenn

$$\varphi(A) = \varphi(A_1) + \varphi(A_2) + \cdots,$$

so sind die Mengen A_n in A meßbar.

Handelt es sich zunächst um zwei Summanden, $A = A_1 + A_2$, so gibt es nach S. 413 eine meßbare Menge (z. B. ein G_δ) $M_1 \geqq A_1$ mit $f(M_1) = \varphi(A_1)$ und ebenso ein $M_2 \geqq A_2$ mit $f(M_2) = \varphi(A_2)$. Setzt man wieder $M = \mathfrak{S}(M_1, M_2)$, $M' = \mathfrak{D}(M_1, M_2)$, so ist

$$\varphi(A) = f(M_1) + f(M_2) = f(M) + f(M') \geqq \varphi(A) + f(M'),$$

also $f(M') = 0$, M' und seine Teilmengen haben das Maß Null. Danach ist $\mathfrak{D}(M_1, A_2) \leqq \mathfrak{D}(M_1, M_2)$ vom Maße 0 und

$$A_1 = \mathfrak{D}(A - A_2, M_1) = \mathfrak{D}\left(A, M_1 - \mathfrak{D}(M_1, A_2)\right)$$

der Durchschnitt von A mit einer meßbaren Menge.

Handelt es sich um beliebig viele Summanden, so folgt aus der vorausgesetzten Gleichung nach den Summenformeln

$$\begin{aligned} \varphi(A) &= \varphi(A_1) + \varphi(A_2) + \varphi(A_3) + \dots \\ &\geqq \varphi(A_1) + \varphi(A_2 + A_3 + \dots) \geqq \varphi(A), \end{aligned}$$

also $\varphi(A) = \varphi(A_1) + \varphi(A - A_1)$, demnach die relative Meßbarkeit von A_1 und jedem A_n.

Hiernach zieht die Gleichung $\varphi(A) = \sum \varphi(A_n)$ die Gleichung $\psi(A) = \sum \psi(A_n)$ nach sich, während das Umgekehrte nicht gilt (vgl. S. 419).

Um nachträglich auch die unbeschränkten Mengen (die man auch von vornherein durch geeignete Modifikationen hätte berücksichtigen können) in den Kreis der Betrachtung zu ziehen, definieren wir für eine beliebige Menge A die Zahlen $\varphi_0(A)$, $\psi_0(A)$, $\varphi(A)$, $\psi(A)$ als obere Schranken der Zahlen $\varphi_0(B)$, $\psi_0(B)$, $\varphi(B)$, $\psi(B)$ für alle in A enthaltenen beschränkten Mengen B, vorausgesetzt, daß diese oberen Schranken existieren (d. h. endlich sind). Wird die Ebene als Summe $E = \mathfrak{S} E_n$ einer aufsteigenden Folge beschränkter Mengen E_n dargestellt, derart, daß zu jeder beschränkten Menge ein sie einschließendes E_n existiert (z. B. als Summe konzentrischer Kreisflächen mit den Radien n), so ist, mit $A_n = \mathfrak{D}(A, E_n)$, $A = \mathfrak{S} A_n$ und man sieht leicht, daß die Gleichungen

$$\varphi_0(A) = \lim \varphi_0(A_n)$$

usw. bestehen, vorausgesetzt, daß die Grenzwerte rechterhand existieren. Alles Bisherige, Symmetrieformeln, Summenformeln usw., überträgt sich damit fast ohne Änderung auf diejenigen unbeschränkten Mengen, deren äußere Inhalte resp. Maße existieren.

Wir bemerken noch, daß manche Theoreme über das Maß von Punktmengen vielleicht ein vertrauteres Gesicht zeigen, wenn man sie in der Sprache der Wahrscheinlichkeitsrechnung ausdrückt. Wenn zwei Mengen P und M meßbar, M insbesondere von positivem Maß ist, so kann man, für $P \leqq M$, den Quotienten $f(P) : f(M)$ oder,

im allgemeinen Falle, den Quotienten $f(\mathfrak{D}(P, M)) : f(M)$ als Wahrscheinlichkeit dafür definieren, daß ein Punkt von M der Menge P angehöre. Betrachten wir nur Teilmengen einer festen Menge M, die das Maß 1 hat, so ist $f(P) = p$ die Wahrscheinlichkeit, daß ein Punkt der Menge P angehöre. Ist $P = \mathfrak{S}(P_1, P_2)$, $P' = \mathfrak{D}(P_1, P_2)$, so ist $p + p' = p_1 + p_2$; p ist die Wahrscheinlichkeit, daß ein Punkt zu P_1 oder P_2 gehöre, p' die Wahrscheinlichkeit, daß ein Punkt sowohl zu P_1 als zu P_2 gehöre. Wenn P_1, P_2 keinen Punkt gemein haben, so ist $p = p_1 + p_2$. Ferner ist $p' = p_1 \cdot \frac{p'}{p_1}$ (für $p_1 > 0$); die Wahrscheinlichkeit, daß ein Punkt gleichzeitig zu P_1 und P_2 gehöre, ist die Wahrscheinlichkeit, daß er zu P_1 gehöre, multipliziert mit der, daß ein Punkt von P_1 zu P_2 gehöre. Die für sogenannte „unabhängige" Ereignisse gültige Formel $p' = p_1 p_2$ ist natürlich hier im allgemeinen nicht richtig. Auch versteht es sich von selbst, daß bei dieser (im ganzen überhaupt willkürlichen) Definition die Wahrscheinlichkeit 0 nicht der Ausdruck der Unmöglichkeit und die Wahrscheinlichkeit 1 nicht der der Gewißheit ist; denn 0 ist die Wahrscheinlichkeit, daß ein Punkt einer Menge vom Maße Null (die noch von der Mächtigkeit des Kontinuums sein kann) angehöre.

§ 4. Beispiele und Anwendungen.

I. Lineare Gebiete und abgeschlossene Mengen. Das Maß eines linearen beschränkten Gebietes $G > 0$ ist sehr einfach zu bestimmen; G ist eine Summe (S. 330) von endlich oder abzählbar vielen zusammenhängenden Gebieten, d. h. offenen Intervallen $(a < x < b)$, und bezeichnet $\delta = b - a$ die Länge eines solchen Intervalls $\Delta = [a, b]$, so ist

$$G = \Delta_1 + \Delta_2 + \cdots, \quad f(G) = \delta_1 + \delta_2 + \cdots.$$

Das Maß einer beschränkten abgeschlossenen Menge F erhält man durch Komplementbildung; man schließe F in ein offenes Intervall Δ ein, dann ist $\Delta - F = G$ ein Gebiet und $f(F) = \delta - f(G)$.

Stellt man $G = \mathfrak{S}(\Delta_1, \Delta_2, \ldots)$ als Summe von Intervallen dar, die nicht paarweise fremd sind, so ist $f(G) \leqq \delta_1 + \delta_2 + \cdots$.

Wir wissen bereits, daß ein z. B. in der Einheitsstrecke $\Delta(0 < x < 1)$ dichtes Gebiet beliebig kleines Maß haben kann; ist $R = \{r_1, r_2, \ldots\}$ eine abzählbare in Δ dichte Menge, etwa die der rationalen Zahlen zwischen 0 und 1, und wählt man zu r_n ein einschließendes Intervall $\Delta_n \leqq \Delta$, so hat das in Δ dichte Gebiet $G = \mathfrak{S}\Delta_n$ ein Maß $f(G) \leqq \Sigma \delta_n$, und diese Summe kann man konvergent und beliebig klein wählen. Ist $D(0 \leqq x \leqq 1)$ die abgeschlossene Einheitsstrecke, so ist die abgeschlossene Menge

$F = D - G$ in D nirgendsdicht und hat doch ein der Zahl 1 beliebig nahe kommendes Maß. Das Maß $f(G) = \psi_0(G)$ ist der innere Inhalt von G, das Maß $f(F) = \varphi_0(F)$ der äußere Inhalt von F, während der äußere Inhalt $\varphi_0(G) = 1$ und der innere Inhalt $\psi_0(F) = 0$ ist.

Die Cantorsche nirgendsdichte perfekte Menge F, bestehend aus denjenigen Zahlen in D, die eine triadische Entwicklung ohne Einsen haben (S. 254), hat das Maß Null; das Gebiet $G = D - F$ ist nämlich die Summe paarweise fremder Intervalle, deren eins die Länge $\frac{1}{3}$, zwei die Länge $\frac{1}{9}$, vier die Länge $\frac{1}{27}$ usw. haben, und hat das Maß $\frac{1}{3}\{1 + \frac{2}{3} + (\frac{2}{3})^2 + \ldots\} = 1$. Man kann die Konstruktion leicht so abändern, daß F ein positives Maß erhält. Ist z. B.

$$G = \left[\frac{1}{3}, \frac{2}{3}\right] + \left[\frac{4}{27}, \frac{5}{27}\right] + \left[\frac{22}{27}, \frac{23}{27}\right] + \cdots,$$

d. h. entsteht G so, daß man das mittlere Drittel der Einheitsstrecke, von den verbleibenden Strecken die mittleren Neuntel, dann die mittleren 27tel nimmt usw., so ist

$$f(G) = \frac{1}{3} + \frac{2}{3}\,\frac{1}{9} + \frac{2 \cdot 8}{3 \cdot 9}\,\frac{1}{27} + \frac{2 \cdot 8 \cdot 26}{3 \cdot 9 \cdot 27}\,\frac{1}{81} + \cdots < \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \cdots = \frac{1}{2};$$

G ist in D dicht und die verbleibende nirgendsdichte perfekte Menge $F = D - G$ hat ein Maß $> \frac{1}{2}$.

Zu der Cantorschen perfekten Menge F vom Maße Null ist noch zu bemerken: alle Teilmengen von F haben natürlich auch das Maß Null, und da F von der Mächtigkeit $\aleph$ des Kontinuums ist, so ist das System dieser Teilmengen von der Mächtigkeit $2^{\aleph}$. Es gibt also ebensoviele ($2^{\aleph}$) meßbare Mengen wie Mengen überhaupt.

Andererseits gibt es auch $2^{\aleph}$ unmeßbare Mengen. Wir hatten (S. 402) die Einheitsstrecke A derart in

$$A = \Sigma A_m = \ldots + A_{-1} + A_0 + A_1 + \cdots$$

gespalten, daß die A_m paarweise zerlegungsgleich (bei Aufwickelung von A auf eine Kreisperipherie unmittelbar kongruent) waren. Es ist leicht einzusehen, daß alle diese Mengen dasselbe äußere und dasselbe innere Maß haben müssen; z. B. war $A_0 = B_0 + C_0$, $A_1 = B_1 + C_1$, B_0 mit B_1 und C_0 mit C_1 kongruent, überdies waren B_0 und C_0 in A_0 relativ meßbar, nämlich Durchschnitte von A_0 mit (einseitig begrenzten) Intervallen, ebenso B_1 und C_1 in A_1 meßbar. Demnach ist

$$\varphi(A_1) = \varphi(B_1) + \varphi(C_1) = \varphi(B_0) + \varphi(C_0) = \varphi(A_0),$$

ebenso $\psi(A_1) = \psi(A_0)$. Die Mengen A_m sind unmeßbar; wegen $1 = f(A) \geqq \Sigma\,\psi(A_m)$ ist ihr inneres Maß Null; das äußere ist positiv, da sonst $f(A) \leqq \Sigma\,\varphi(A_m)$ ebenfalls verschwinden würde. Eine dieser Mengen, A_0, erhielt man, indem man aus $\aleph$ paarweise fremden,

abzählbaren Mengen, den damaligen P_x, je einen Punkt auswählte, und diese Wahl ist auf $\aleph_0^{\aleph} = 2^{\aleph}$ Weisen möglich.

Auch die Mengen $A_{m+1} + A_{m+2} + \cdots + A_{m+n}$ sind, bei festgehaltener natürlicher Zahl n, paarweise in derselben Art wie oben zerlegungsgleich (auf dem Kreise kongruent) und haben dasselbe äußere wie dasselbe innere Maß; das letztere ergibt sich wieder $= 0$. Es ist also

$$\psi(A_1 + A_2 + \dots + A_n) = \psi(A_1) + \psi(A_2) + \dots + \psi(A_n),$$

während die Gleichung

$$\varphi(A_1 + A_2 + \dots + A_n) = \varphi(A_1) + \varphi(A_2) + \dots + \varphi(A_n)$$

jedenfalls von einem gewissen n ab nicht mehr richtig sein kann, da die linke Seite $\leqq 1$ bleibt (vgl. S. 416).

Setzt man noch

$$P_n = A_0 + A_1 + A_{-1} + \dots + A_n + A_{-n},$$

so ist $A = \mathfrak{S}(P_1, P_2, \dots)$ Summe einer aufsteigenden Mengenfolge und demnach konvergiert $\varphi(P_n)$ nach $\varphi(A) = 1$, während $\psi(P_n) = 0$ nicht nach $\psi(A) = 1$ konvergiert. Entsprechend kann man eine absteigende Folge von Mengen bilden, deren äußeres Maß nicht nach dem des Durchschnitts konvergiert (vgl. S. 412).

Auf Grund der drei ersten Lebesgueschen Postulate (S. 401) wäre den Mengen A_m der Inhalt Null zuzuschreiben; man sieht, daß die Tragweite dieser Forderungen noch durchaus nicht erschöpft ist und die Frage nach einer etwaigen, noch weitere Mengen umfassenden Inhaltsbestimmung offen bleibt.

II. Mengen dyadischer Brüche. Betrachten wir eine irrationale[1] Zahl x zwischen 0 und 1 und entwickeln sie in einen dyadischen Bruch:

$$x = \frac{x_1}{2} + \frac{x_2}{2^2} + \frac{x_3}{2^3} + \dots = (x_1, x_2, x_3, \dots) \qquad (x_n = 0, 1).$$

Unter den n ersten Ziffern mögen sich p Nullen und $q = n - p$ Einsen befinden. Dann gilt der Satz (E. Borel):

Die Menge der x, für die $\lim \frac{p}{n} = \frac{1}{2}$, hat das Maß 1.

Oder: das Komplement, d. h. die Menge derjenigen x, für die $\frac{p}{n}$ entweder überhaupt nicht oder nicht nach $\frac{1}{2}$ konvergiert, hat das Maß 0. Es besteht also die Wahrscheinlichkeit 1 dafür, daß

[1] Wir beschränken uns auf diese wegen der Eindeutigkeit der dyadischen Entwicklung; natürlich würde es auch genügen, die dyadisch rationalen Zahlen auszuschließen oder unter deren beiden dyadischen Entwicklungen eine zu bevorzugen.

die dyadische Entwicklung von x asymptotisch ebensoviele Nullen wie Einsen aufweise.

Dieser Satz ist merkwürdig. Auf der einen Seite erscheint er als plausible Übertragung des „Gesetzes der großen Zahlen" ins Unendliche; andererseits ist doch die Existenz eines Limes für eine Zahlenfolge, noch dazu eines vorgeschriebenen Limes, ein sehr spezieller Fall, den man a priori für äußerst unwahrscheinlich halten sollte.

Um den Satz zu beweisen, bemerken wir zunächst: die Menge der x, deren Entwicklung mit n gegebenen Ziffern $a_1, a_2, \ldots, a_n$ beginnt, ist offenbar die Menge der irrationalen Punkte der Strecke

$$\frac{a_1}{2} + \cdots + \frac{a_n}{2^n} < x < \frac{a_1}{2} + \cdots + \frac{a_n + 1}{2^n},$$

hat also das Maß $\frac{1}{2^n}$ (die Menge der rationalen Punkte kommt, als vom Maße 0, nicht in Betracht). Die Menge der x, für die unter den n ersten Ziffern p Nullen und q Einsen vorhanden sind, setzt sich aus $\binom{n}{p} = \frac{n!}{p!\,q!}$ solchen paarweise fremden Strecken zusammen und hat also das Maß $\binom{n}{p}\frac{1}{2^n}$.

Ferner sei ε eine positive Zahl und $M(n, \varepsilon)$ die Menge derjenigen x, für die $\left|\frac{p}{n} - \frac{1}{2}\right| \geqq \varepsilon$. Das Maß dieser Menge, die sich immer noch aus einer endlichen Anzahl der zuvor betrachteten Mengen zusammensetzt, ist

$$m(n, \varepsilon) = \sum^{*} \binom{n}{p} \frac{1}{2^n},$$

die Summe $\sum^{*}$ über diejenigen p erstreckt, wo $\left|\frac{p}{n} - \frac{1}{2}\right| \geqq \varepsilon$.

Um diese Zahl elementar abzuschätzen, wenden wir auf die für beliebige x, y geltende Identität

$$\sum \binom{n}{p} x^p y^q = (x + y)^n$$

(wo die Summe über alle $p = 0, 1, \ldots, n$ zu erstrecken ist und $q = n - p$), wiederholt den Differentiationsprozeß $x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}$ an. Setzen wir

$$u = x + y, \quad v = x - y,$$

so folgt

$$\sum \binom{n}{p} x^p y^q = u^n,$$

$$\sum \binom{n}{p} (p - q) x^p y^q = n u^{n-1} v,$$

$$\sum \binom{n}{p} (p - q)^2 x^p y^q = n u^n + n(n-1) u^{n-2} v^2,$$

$$\Sigma \binom{n}{p} (p-q)^3 x^p y^q = n(3n-2)u^{n-1}v + n(n-1)(n-2)u^{n-3}v^3,$$
$$\Sigma \binom{n}{p} (p-q)^4 x^p y^q = n(3n-2)u^n + n(n-1)(6n-8)u^{n-2}v^2 + n(n-1)(n-2)(n-3)u^{n-4}v^4.$$

Also für $x = y = \frac{1}{2}$

$$\Sigma \binom{n}{p} (p-q)^4 \frac{1}{2^n} = 3n^2 - 2n < 3n^2$$

oder

$$\Sigma \binom{n}{p} \frac{1}{2^n} \left(\frac{p}{n} - \frac{1}{2}\right)^4 < \frac{3}{16} \frac{1}{n^2}.$$

Hieraus folgt

$$\frac{3}{16} \frac{1}{n^2} > \overset{*}{\Sigma} \binom{n}{p} \frac{1}{2^n} \left(\frac{p}{n} - \frac{1}{2}\right)^4 \geqq \varepsilon^4 \cdot \overset{*}{\Sigma} \binom{n}{p} \frac{1}{2^n},$$

$$m(n, \varepsilon) < \frac{3}{16\varepsilon^4} \cdot \frac{1}{n^2}.$$

Sodann sei $M(\varepsilon) = \operatorname{Lim\,sup}_n M(n, \varepsilon)$ die Menge der x, für die unendlich oft $\left|\frac{p}{n} - \frac{1}{2}\right| \geqq \varepsilon$. Es ist also

$$M(\varepsilon) = \mathfrak{D}\left(S(1, \varepsilon),\ S(2, \varepsilon), \ldots\right),$$
$$S(n, \varepsilon) = \mathfrak{S}\left(M(n, \varepsilon),\ M(n+1, \varepsilon), \ldots\right).$$

Das Maß von $S(n, \varepsilon)$ ist

$$s(n, \varepsilon) \leqq m(n, \varepsilon) + m(n+1, \varepsilon) + \ldots < \frac{3}{16\varepsilon^4} \left(\frac{1}{n^2} + \frac{1}{(n+1)^2} + \ldots\right);$$

wegen der Konvergenz der Reihe $\Sigma \frac{1}{n^2}$ ist daher das Maß von $M(\varepsilon)$

$$m(\varepsilon) = \lim s(n, \varepsilon) = 0.$$

Die Menge $M(\varepsilon)$ hat also das Maß Null, folglich auch die Menge

$$M = \underset{\varepsilon}{\mathfrak{S}} M(\varepsilon) = \mathfrak{S}\left(M(1),\ M(\tfrac{1}{2}),\ M(\tfrac{1}{3}), \ldots\right)$$

und dies ist gerade die Menge derjenigen x, für die nicht $\lim \frac{p}{n} = \frac{1}{2}$ (sei es, daß der Limes nicht existiert oder nicht $\frac{1}{2}$ ist), d. h. zu denen sich eine positive Zahl ε angeben läßt derart, daß unendlich oft $\left|\frac{p}{n} - \frac{1}{2}\right| \geqq \varepsilon$. Das Komplement von M hat also das Maß 1, womit der Borelsche Satz bewiesen ist.

Man kann durch weiteren Fortschritt in der Kette der obigen Identitäten oder auch durch die in der Wahrscheinlichkeitsrechnung übliche Herbeiziehung von Integralen sogar noch beweisen, daß die Konvergenz von $\frac{p}{n}$ an $\frac{1}{2}$ mit der Wahrscheinlichkeit 1 eine gewisse Stärke hat, z. B. daß für jede Zahl $\vartheta < \frac{1}{2}$ die Menge der Punkte, für die $\lim \left(\frac{p}{n} - \frac{1}{2}\right) \cdot n^\vartheta = 0$, das Maß 1 hat.

Es ist evident, daß ein dem Borelschen Satz analoger für triadische Brüche usw. gilt; z. B. hat die Menge der Dezimalbrüche zwischen 0 und 1, die asymptotisch jede Ziffer 0, 1, ..., 9 gleich oft enthalten, das Maß 1. Die Menge der Dezimalbrüche, in denen eine Ziffer fehlt, hat das Maß Null; so auch die Cantorsche perfekte Menge der triadischen Brüche, in denen die Ziffer 1 fehlt.

III. Mengen von Kettenbrüchen. Jede irrationale Zahl x zwischen 0 und 1 läßt sich auf eine und nur eine Weise in einen Kettenbruch

$$x = 1 : x_1 + 1 : x_2 + 1 : x_3 + \ldots = (x_1, x_2, x_3, \ldots)$$

entwickeln, wo die „Teilnenner" x_n natürliche Zahlen sind.

Wir wollen das Maß der Menge derjenigen Zahlen x ermitteln oder wenigstens abschätzen, deren Teilnenner durch Ungleichungen der Form

$$a_n \leqq x_n < b_n \quad \text{oder} \quad a_n \leqq x_n$$

eingeengt sind, wobei die a_n natürliche Zahlen, die b_n natürliche Zahlen $> a_n$ (also jedenfalls $\geqq 2$) sind.

Bezeichnen wir den Komplex der ersten n Ziffern des Kettenbruchs mit $\xi_n = (x_1, x_2, \ldots, x_n)$; zugleich soll dies den Wert des endlichen Kettenbruchs

$$\xi_n = 1 : x_1 + 1 : x_2 + 1 : \ldots : x_{n-1} + 1 : x_n$$

bedeuten. Dieser n^{te} Näherungsbruch ist bekanntlich $\xi_n = \frac{u_n}{v_n}$, wobei

$$u_1 = 1\ , \ u_2 = x_2 \qquad , \quad u_n = x_n u_{n-1} + u_{n-2},$$
$$v_1 = x_1, \ v_2 = x_2 x_1 + 1, \quad v_n = x_n v_{n-1} + v_{n-2};$$

um die für $n > 2$ gültigen Rekursionsformeln auch für $n = 1, 2$ aufrechtzuerhalten, ist noch

$$u_{-1} = 1, \quad u_0 = 0,$$
$$v_{-1} = 0, \quad v_0 = 1$$

zu setzen. Allgemein ist $u_{n-1} v_n - u_n v_{n-1} = (-1)^n$.

Die Menge $M(x_1, x_2, \ldots, x_n) = M(\xi_n)$ der x, die mit gegebenen n Ziffern beginnen, ist die Menge der irrationalen Zahlen des Intervalls mit den Endpunkten

$$(x_1, x_2, \ldots, x_n) = \xi_n = \frac{u_n}{v_n}$$

und

$$(x_1, x_2, \ldots, x_n, 1) = (\xi_n, 1) = \frac{u_n + u_{n-1}}{v_n + v_{n-1}}.$$

Die Länge dieses Intervalls oder das Maß jener Menge ist

$$m(\xi_n) = \left| \frac{u_n}{v_n} - \frac{u_n + u_{n-1}}{v_n + v_{n-1}} \right| = \frac{1}{v_n (v_n + v_{n-1})} = \frac{1}{v_{n-1}} \left(\frac{1}{v_n} - \frac{1}{v_n + v_{n-1}} \right)$$

oder, wenn wir die Abhängigkeit von der letzten Ziffer x_n hervorheben,

$$m(\xi_n) = m(\xi_{n-1}, x_n) = \frac{1}{v_{n-1}}\left(\frac{1}{x_n v_{n-1} + v_{n-2}} - \frac{1}{(x_n+1)v_{n-1} + v_{n-2}}\right).$$

Aus solchen Intervallen (von den rationalen Zahlen abgesehen), die für bestimmtes n wegen der Eindeutigkeit der Kettenbruchdarstellung paarweise fremd sind, setzen sich die im folgenden auftretenden Mengen als endliche oder abzählbare Summen zusammen.

Wir halten jetzt nur die ersten $n-1$ Ziffern $\xi_{n-1} = (x_1, x_2, \ldots, x_{n-1})$ fest,[1] während wir x_n durch die Bedingung $a \leqq x_n < b$ ($a < b$ natürliche Zahlen) fixieren. Die Menge dieser Zahlen x sei

$$P(\xi_{n-1}) = M(\xi_{n-1}, a) + M(\xi_{n-1}, a+1) + \ldots + M(\xi_{n-1}, b-1);$$

ihr Maß ist nach der obigen Formel

$$p(\xi_{n-1}) = \frac{1}{v_{n-1}}\left(\frac{1}{a v_{n-1} + v_{n-2}} - \frac{1}{b v_{n-1} + v_{n-2}}\right).$$

Falls die Ziffer x_n nur einseitig durch $a \leqq x_n$ begrenzt wird, erhalten wir

$$P(\xi_{n-1}) = M(\xi_{n-1}, a) + M(\xi_{n-1}, a+1) + \ldots$$

und

$$p(\xi_{n-1}) = \frac{1}{v_{n-1}} \cdot \frac{1}{a v_{n-1} + v_{n-2}};$$

eine besondere Berücksichtigung dieses Falles ist entbehrlich, wenn man in bekannter Weise den uneigentlichen Wert $b = \infty$ zuläßt. Ist endlich noch $a = 1$, so erhält man wieder die Menge $M(\xi_{n-1})$ aller mit ξ_{n-1} beginnenden Kettenbrüche, mit dem Maß

$$m(\xi_{n-1}) = \frac{1}{v_{n-1}} \cdot \frac{1}{v_{n-1} + v_{n-2}}.$$

Daraus folgt

$$\frac{p(\xi_{n-1})}{m(\xi_{n-1})} = \frac{v_{n-1} + v_{n-2}}{a v_{n-1} + v_{n-2}} - \frac{v_{n-1} + v_{n-2}}{b v_{n-1} + v_{n-2}}$$

oder mit

$$\vartheta = \frac{v_{n-2}}{v_{n-1}}:$$

$$\frac{p(\xi_{n-1})}{m(\xi_{n-1})} = \frac{1+\vartheta}{a+\vartheta} - \frac{1+\vartheta}{b+\vartheta} = \varphi(a, b, \vartheta).$$

Nun liegt ϑ zwischen 0 und 1, da die Zahlen v_n mit dem Index wachsen; bezeichnet man also von der Funktion $\varphi(a, b, \vartheta)$ das Minimum und Maximum im Intervall $0 \leqq \vartheta \leqq 1$ mit

$$\varrho = \varrho(a, b), \quad \sigma = \sigma(a, b),$$

wobei der Bedeutung dieser Funktion nach $0 < \varrho \leqq \sigma \leqq 1$, so ist

$$\varrho \cdot m(\xi_{n-1}) \leqq p(\xi_{n-1}) \leqq \sigma \cdot m(\xi_{n-1}).$$

---

[1] Wie dies und das folgende für $n = 1$ aufzufassen ist, liegt auf der Hand.

Nun lassen wir auch die ersten $n-1$ Ziffern noch variieren, während die Bedingung $a \leqq x_n < b$ für x_n festgehalten wird; d. h. wir bilden die Menge

$$M = M(\xi_{n-1}) + M(\eta_{n-1}) + \cdots$$

der Kettenbrüche, deren Anfangsziffernkomplex ξ_{n-1} eine irgendwie vorgeschriebene (natürlich höchstens abzählbare) Menge $\{\xi_{n-1}, \eta_{n-1}, \ldots\}$ durchläuft, während alle folgenden Ziffern frei variieren, und die Menge

$$P = P(\xi_{n-1}) + P(\eta_{n-1}) + \cdots$$

derjenigen Kettenbrüche von M, deren n^{te} Ziffer außerdem noch an $a \leqq x_n < b$ gebunden ist. Da die Maße m, p dieser Mengen die Summen der Maße der entsprechenden Summanden sind, so gilt

$$\varrho m \leqq p \leqq \sigma m, \quad \varrho \leqq \frac{p}{m} \leqq \sigma.$$

Endlich sei nun M_n die Menge der Kettenbrüche, für die gleichzeitig

$$a_1 \leqq x_1 < b_1, \quad a_2 \leqq x_2 < b_2, \ldots, \quad a_n \leqq x_n < b_n,$$

während die übrigen Ziffern frei variieren. Dann steht M_n zu M_{n-1} in eben dem Verhältnis wie zuvor P zu M, falls man a, b durch a_n, b_n ersetzt, und für die Maße m_n findet man

$$\varrho_n \leqq \frac{m_n}{m_{n-1}} \leqq \sigma_n,$$

$$\varrho_n = \varrho(a_n, b_n), \quad \sigma_n = \sigma(a_n, b_n).$$

Dabei ist unter M_0 die Menge aller Kettenbrüche zu verstehen, also $m_0 = 1$. Durch Multiplikation findet man

$$\varrho_1 \varrho_2 \cdots \varrho_n \leqq m_n \leqq \sigma_1 \sigma_2 \ldots \sigma_n.$$

Die Menge $M = \mathfrak{D}(M_1, M_2, \ldots)$ der Kettenbrüche, die für jedes n der Bedingung $a_n \leqq x_n < b_n$ genügen, hat als Durchschnitt einer absteigenden Folge das Maß $m = \lim m_n$. Andererseits konvergieren auch die Produkte $\varrho_1 \varrho_2 \cdots \varrho_n$, die wegen $0 < \varrho_n \leqq 1$ eine absteigende Folge bilden, gegen einen gewissen Grenzwert $\varrho_1 \varrho_2 \ldots$, ebenso die $\sigma_1 \sigma_2 \ldots \sigma_n$ gegen $\sigma_1 \sigma_2 \ldots$, und man hat endlich

$$\varrho_1 \varrho_2 \cdots \leqq m \leqq \sigma_1 \sigma_2 \cdots.$$

Es kommt nur noch auf die Bestimmung der ϱ und σ an. Die Ableitung der fraglichen Funktion $\varphi(\alpha, \beta, \vartheta)$ nach ϑ hat das Zeichen von

$$(a-1)(b-1) - (1+\vartheta)^2,$$

ist also für $a = 1$ immer negativ, für $a > 1$ immer $\geqq 0$ bis auf die beiden „singulären“ Fälle $a = 2$, $b = 3$ und $a = 2$, $b = 4$, wo sie inner-

halb des Intervalls das Zeichen wechselt. Abgesehen von den singulären Fällen, wo man

$$\varrho(2,3)=\tfrac{1}{6},\quad \sigma(2,3)=3-2\sqrt{2},$$
$$\varrho(2,4)=\tfrac{1}{4},\quad \sigma(2,4)=2-\sqrt{3}$$

findet, tritt also Maximum und Minimum der Funktion an den Intervallgrenzen ein und es ist für $a=1$

$$\varrho(1,b)=\frac{b-1}{b+1},\quad \sigma(1,b)=\frac{b-1}{b},$$

für $a>1$

$$\varrho(a,b)=\frac{1}{a}-\frac{1}{b},\quad \sigma(a,b)=\frac{2}{a+1}-\frac{2}{b+1};$$

insbesondere noch für $b=\infty$ (wenn also die Ungleichung $a\leqq x$ an Stelle von $a\leqq x<b$ tritt) und $a\geqq 1$

$$\varrho(a,\infty)=\frac{1}{a},\quad \sigma(a,\infty)=\frac{2}{a+1}.$$

Wir fragen nun, in welchem Fall die Menge M positives Maß hat; dazu ist notwendig, daß das Produkt $\sigma_1\sigma_2\ldots$ positiv sei, also die Reihe $\Sigma(1-\sigma_n)$ konvergiere, und hinreichend, daß $\varrho_1\varrho_2\ldots$ positiv sei, also die Reihe $\Sigma(1-\varrho_n)$ konvergiere. Nun ist für $a>1$ im allgemeinen Fall

$$1-\sigma(a,b)=\frac{a-1}{a+1}+\frac{2}{b+1}\geqq\frac{1}{3}$$

und in den singulären $1-\sigma$ eine feste positive Zahl; diese Fälle dürfen also jedenfalls, wenn $\Sigma(1-\sigma_n)$ konvergieren soll, nicht unendlich oft auftreten, es müssen fast alle $a_n=1$ sein. Für $a=1$ ist $1-\sigma(1,b)=\frac{1}{b}$, und danach ist noch die Konvergenz von $\Sigma\frac{1}{b_n}$ notwendig. Da dann $1-\varrho(1,b)=\frac{2}{b+1}$ und mit $\Sigma\frac{1}{b_n}$ auch $\Sigma\frac{1}{b_n+1}$ konvergiert, ist die gefundene Bedingung auch hinreichend. Also:

Die Menge M der durch die Ungleichungen $a_n\leqq x_n<b_n$ eingeschränkten Kettenbrüche hat dann und nur dann positives Maß, wenn fast alle $a_n=1$ sind und die Reihe $\Sigma\frac{1}{b_n}$ konvergiert. Falls an Stelle einer solchen Ungleichung die einseitige $a_n\leqq x_n$ tritt, ist $\frac{1}{b_n}=0$ zu setzen, während die Ungleichung $x_n<b_n$ soviel wie $a_n=1$ besagt. Nimmt man alle $a_n=1$, so hat die durch die Ungleichungen $x_n<b_n$ definierte Kettenbruchmenge M ein Maß, das zwischen

$$\Pi\varrho_n=\Pi\frac{b_n-1}{b_n+1}\ \text{und}\ \Pi\sigma_n=\Pi\frac{b_n-1}{b_n}$$

liegt, und das dann und nur dann positiv ist, wenn die Reihe $\Sigma \frac{1}{b_n}$ konvergiert.

Die Bedeutung des obigen Satzes kann man etwa so bezeichnen, daß ein zu langsames oder zu schnelles Wachstum der Teilnenner x_n die Wahrscheinlichkeit Null erhält. Setzt man z. B. alle $b_n = b$, alle $a_n = 1$, so hat die Menge der Kettenbrüche, deren sämtliche Teilnenner $< b$, das Maß 0; und die Summe dieser Mengen für $b = 2, 3, 4, \ldots$ oder die Menge aller Kettenbrüche mit beschränkter Ziffernfolge hat ebenfalls das Maß 0. Dies bleibt bestehen, wenn man nur für eine bestimmte Teilfolge natürlicher Zahlen p die $b_p = b$, alle übrigen $= \infty$ setzt; die Kettenbrüche, bei denen eine (im voraus fixierte) Teilfolge der x_n beschränkt ist, bilden ebenfalls eine Menge vom Maße Null. — Andererseits, setzt man $a_n = a_{n+1} = \ldots = 2$, $a_1 = \ldots = a_{n-1} = 1$, und alle $b_n = \infty$, so hat die Menge der Kettenbrüche, wo $x_n, x_{n+1}, \ldots \geqq 2$, das Maß 0, und die Summe dieser Mengen für $n = 1, 2, 3, \ldots$ oder die Menge der Kettenbrüche mit nur endlich vielen Einsen hat auch noch das Maß 0: es besteht also die Wahrscheinlichkeit 1 für das Auftreten unendlich vieler Einsen unter den Teilnennern (während das Auftreten von lauter Einsen in einer im voraus fixierten Teilfolge die Wahrscheinlichkeit 0 hatte). Dagegen hat z. B. die Menge der Kettenbrüche, wo $x_n < kn^2$ (k eine natürliche Zahl > 1), wegen der Konvergenz der Reihe $\Sigma \frac{1}{n^2}$ ein positives Maß, das zwischen den Grenzen liegt

$$\Pi \frac{kn^2-1}{kn^2+1} = \Pi\left(1 - \frac{1}{kn^2}\right) : \Pi\left(1 + \frac{1}{kn^2}\right) = \sin\frac{\pi}{\sqrt{k}} : \operatorname{sh}\frac{\pi}{\sqrt{k}}$$

und

$$\Pi \frac{kn^2-1}{kn^2} = \Pi\left(1 - \frac{1}{kn^2}\right) = \sin\frac{\pi}{\sqrt{k}} : \frac{\pi}{\sqrt{k}},$$

z. B. für $k = 4$ zwischen den Grenzen $1 : \operatorname{sh}\frac{\pi}{2} = 2 : \left(e^{\frac{\pi}{2}} - e^{-\frac{\pi}{2}}\right)$ und $2 : \pi$ (0,4345 ... und 0,6366 ...).

IV. Meßbare Funktionen. In der (beschränkten) euklidischen Punktmenge A sei eine reelle Funktion $f(x)$ definiert. Damit sind für jedes reelle y die Mengen

$$A(y) = \text{Menge der Stellen } x, \text{ wo } f(x) \geqq y,$$
$$A(y+0) = \text{Menge der Stellen } x, \text{ wo } f(x) > y$$

definiert, die wir schon in Kap. I, § 11 und IX, § 5 behandelt haben. Mit Lebesgue nennen wir die Funktion $f(x)$ meßbar, wenn die Menge $A(y)$ für jedes y meßbar ist. Es ist dann auch die ganze Menge

$$A = \mathfrak{S}_n A(-n) \qquad (n = 1, 2, 3, \ldots)$$

und jede Menge $$A(y+0) = \mathfrak{S}_n A\left(y+\frac{1}{n}\right)$$
meßbar (in einer nicht meßbaren Menge gibt es keine meßbaren Funktionen). Bezeichnet man also die meßbaren Teilmengen von A mit M, so sind nach unserer damaligen Ausdrucksweise die meßbaren Funktionen in A von der Klasse (M, M), wo die M ein $(\sigma\delta)$-System bilden und ihre eigenen Komplemente sind, d. h. $A-M$ wieder ein M ist; daraus folgt, daß Summe und Differenz von zwei meßbaren Funktionen und der Limes einer konvergenten Folge meßbarer Funktionen wieder eine meßbare Funktion ist. Mit $f(x)$ ist auch ihr Betrag $|f(x)|$, ihr Quadrat $f(x)^2$, ihr Produkt $c\cdot f(x)$ mit einer Konstanten und, falls $f(x)$ in A nirgends verschwindet, ihr reziproker Wert $1:f(x)$ meßbar; das Produkt von zwei meßbaren Funktionen ist meßbar, da mit f und g auch $f\pm g$, $(f+g)^2-(f-g)^2 = 4fg$ und fg meßbar sind. Nehmen wir A als meßbar, so ist eine in A stetige oder oberhalb stetige Funktion meßbar, denn die Mengen $A(y)$ sind dann in A abgeschlossen oder Durchschnitte von A mit abgeschlossenen, d. h. meßbaren Mengen (S. 393); ebenso die unterhalb stetigen und alle aus stetigen durch Limesbildung entstehenden Funktionen, d. h. die Baireschen Funktionen. Die vier Ableitungen einer stetigen Funktion sind meßbar (S. 396). Die Meßbarkeit einer Funktion steht im engsten Zusammenhange mit ihrer Integrabilität (§ 5).

Wir knüpfen daran einige Betrachtungen über eine in A konvergente Folge meßbarer Funktionen, deren Limes $f(x) = f_n(x)$ also wieder meßbar ist. Es sei für ein bestimmtes positives ε
$$A_n = \text{Menge der Stellen, wo } |f_n(x)-f(x)| < \varepsilon,$$
$$P_n = \mathfrak{D}(A_n, A_{n+1}, \dots)$$
$$= \text{Menge der Stellen, wo } |f_p(x)-f(x)| < \varepsilon \text{ für } p \geqq n,$$
und
$$B_n = A - A_n,\ Q_n = A - P_n = \mathfrak{S}(B_n, B_{n+1}, \dots).$$
Da nun jeder Punkt x fast allen Mengen A_n angehört, so ist
$$A = \text{Lim inf } A_n = \mathfrak{S}(P_1, P_2, \dots),$$
$$0 = \text{Lim sup } B_n = \mathfrak{D}(Q_1, Q_2, \dots).$$
Diese Mengen sind meßbar; das Maß von Q_n (und erst recht von B_n) konvergiert mit $\frac{1}{n}$ nach Null, das Maß von P_n (und erst recht von A_n) konvergiert nach dem Maß von A.

Deuten wir jetzt auch die Abhängigkeit von ε an, indem wir $A(n, \varepsilon)$ statt A_n schreiben usw. Man kann demgemäß für beliebig vorgeschriebene positive Zahlen ε, σ die natürliche Zahl n so groß wählen, daß $f(Q(n, \varepsilon)) < \sigma$.

Wählen wir eine Folge positiver Zahlen ε_k, die nach 0 konvergieren, und eine Folge positiver Zahlen σ_k mit konvergenter Summe $\sigma = \sigma_1 + \sigma_2 + \ldots$. Für jeden Index k können wir n_k so groß wählen, daß $f(Q(n_k, \varepsilon_k)) < \sigma_k$. Die Menge $Q = \mathfrak{S}_k Q(n_k, \varepsilon_k)$ hat dann ein Maß $< \sigma_1 + \sigma_2 + \ldots = \sigma$, das man demgemäß beliebig klein machen kann. In der Komplementärmenge

$$P = A - Q = \mathfrak{D}_k P(n_k, \varepsilon_k)$$

ist, für jedes k,

$$|f_n(x) - f(x)| < \varepsilon_k \text{ für } n \geqq n_k,$$

d. h. in P ist die Konvergenz gleichmäßig. Jede konvergente Folge meßbarer Funktionen konvergiert gleichmäßig bis auf eine Menge von beliebig kleinem Maß.

Man folgert hieraus, daß sämtliche Baireschen Funktionen bis auf eine Menge von beliebig kleinem Maß stetig sind; d. h. zu jedem positiven σ gibt es eine Menge Q von einem Maße $< \sigma$ derart, daß die auf $P = A - Q$ eingeschränkte Teilfunktion $f(x|P)$ stetig ist.[1] Diese Eigenschaft überträgt sich nämlich von einer konvergenten Funktionenfolge auf ihren Limes. In der Tat wähle man bei gegebenem $\sigma = \sigma_1 + \sigma_2 + \ldots$ die Menge Q_n vom Maße $< \sigma_n$ so, daß $f_n(x|P_n)$ in $P_n = A - Q_n$ stetig ist; die Menge $Q' = \mathfrak{S} Q_n$ hat ein Maß $< \sigma$ und in der Menge $P' = A - Q' = \mathfrak{D} P_n$ sind alle Funktionen $f_n(x|P')$ stetig. Für eine geeignete Menge Q'' vom Maße $f(Q'') < \sigma - f(Q')$ ist die Konvergenz in $P = P' - Q'' = A - (Q' + Q'')$ gleichmäßig, also der Limes $f(x|P)$ stetig, wobei $Q = A - P = Q' + Q''$ vom Maße $< \sigma$ ist.

Betrachten wir noch einen iterierten Limes meßbarer Funktionen

$$f(x) = \lim_p \lim_q f_{pq}(x)$$

oder

$$f(x) = \lim_p f_p(x), \quad f_p(x) = \lim_q f_{pq}(x).$$

Bedeuten $A(p, q, \varepsilon)$, $B(p, q, \varepsilon)$ die Mengen der Stellen, wo

$$|f_{pq}(x) - f(x)| < \varepsilon \text{ resp. } \geqq \varepsilon,$$

so kann man durch passende Wahl der Zahlen p, q das Maß von $B(p, q, \varepsilon)$ beliebig klein machen; man kann nämlich der Menge, wo $|f_p(x) - f(x)| \geqq \frac{1}{2}\varepsilon$, durch passende Wahl von p ein Maß $< \frac{1}{2}\sigma$, und dann der Menge, wo $|f_{pq}(x) - f_p(x)| \geqq \frac{1}{2}\varepsilon$, durch passende Wahl von q ein Maß $< \frac{1}{2}\sigma$ geben, wonach die Summe beider Mengen

[1] Dies besagt natürlich nicht, daß die Gesamtfunktion $f(x|A)$ in den Punkten von P stetig sei.

und die darin enthaltene Menge $B(p, q, \varepsilon)$ ein Maß $< \sigma$ bekommt. Man wähle wieder die Zahlen ε_k nach 0 konvergierend und die Zahlen σ_k mit konvergenter Summe, sodann die Zahlen p_k, q_k derart, daß $B_k = B(p_k, q_k, \varepsilon_k)$ ein Maß $< \sigma_k$ hat; in $A_k = A(p_k, q_k, \varepsilon_k)$ ist $|f_{p_k q_k}(x) - f(x)| < \varepsilon_k$. In der Menge $P = \text{Lim inf } A_k$ ist dann $\lim\limits_k f_{p_k q_k}(x) = f(x)$. Man setze noch

$$Q = A - P = \text{Lim sup } B_k = \mathfrak{D}(Q_1, Q_2, \ldots),$$

$$Q_k = \mathfrak{S}(B_k, B_{k+1}, \ldots).$$

Die Menge Q_k hat ein Maß $< \sigma_k + \sigma_{k+1} + \ldots$, das also mit $\frac{1}{k}$ nach 0 konvergiert, und Q hat demnach das Maß Null. Also: bis auf eine Menge vom Maße Null ist der wiederholte Limes einer Doppelfolge von meßbaren Funktionen $f_{pq}(x)$ auch als Limes einer einfachen Folge von Funktionen $f_{p_k q_k}(x)$ darstellbar.

Hieraus schließt man ähnlich wie vorhin, daß jede Bairesche Funktion bis auf eine Menge vom Maße Null Limes stetiger Funktionen ist; die wiederholte Limesbildung führt, wenn man Mengen vom Maße Null nicht beachtet, nicht weiter als die einfache.

Daß es unmeßbare Funktionen gibt, geht aus der Existenz unmeßbarer Mengen direkt hervor. Die mehrfach (S. 402, 418) erwähnte Spaltung der Einheitsstrecke $A = \Sigma A_m$ in unmeßbare Bestandteile liefert z. B. eine unmeßbare Funktion, wenn man in A_m einfach $f(x) = m$ setzt.

Hiervon kann man mit H. Lebesgue eine interessante Anwendung machen. Das Zermelosche Wohlordnungsverfahren (S. 136) verlangt, jeder nicht verschwindenden Teilmenge von A ein zu ihr gehöriges Element zuzuordnen. Insbesondere muß also jeder der S. 401 betrachteten Teilmengen

$$P_x = \{\ldots, x - \alpha, x, x + \alpha, \ldots\}$$

ein zu ihr gehöriger Punkt

$$f(x) = x - \alpha m(x)$$

entsprechen, wo $m(x)$ eine von x abhängige ganze Zahl ist; wir haben hierbei wieder x von der Beschränkung $0 \leqq x < 1$ befreit, sodaß Werten von x mit ganzzahliger Differenz derselbe Punkt der Einheitsstrecke A entspricht. Da nun allen Punkten von P_x derselbe Punkt zugeordnet sein soll, so darf $f(x)$ bei einer Änderung von x um Vielfache von α nur ganzzahlige Änderungen erfahren, d. h. $f(x + \alpha) - f(x)$ muß eine ganze Zahl sein. Das gibt wegen der Irrationalität von α

$$m(x + \alpha) = m(x) + 1.$$

Die Funktion $m(x)$ nimmt also in jeder Menge P_x alle ganzzahligen Werte an und jeden nur einmal; wir erhalten damit genau eine solche Zerlegung $A = \Sigma A_m$, wie wir sie betrachteten, wobei A_0 die Menge der Stellen in A ist, wo $m(x) = 0$, und A_m die Menge derer, wo $m(x) = m$. Die Funktionen $m(x)$ und $f(x)$ sind also nicht meßbar, und diese Tatsache, daß die Zermelosche Auswahl der ausgezeichneten Elemente auf unmeßbare, also „analytisch nicht darstellbare" Funktionen[1] führt, erklärt in gewissem Sinne die Unzugänglichkeit des Wohlordnungsproblems.

§ 5. Das Lebesguesche Integral.

Wir ordnen einer reellen nichtnegativen Funktion in gewisser Weise eine Punktmenge, ihre „Ordinatenmenge", zu und definieren das Integral der Funktion als Inhalt resp. Maß der Punktmenge; je nachdem erhalten wir das Riemannsche oder das Lebesguesche Integral. Das letztere zeichnet sich durch die weit umfassendere Menge integrabler Funktionen und durch den erleichterten Übergang von einer Funktionenfolge zu ihrem Limes aus: wir wollen unsere Betrachtung auf eine Einführung in diesen modernen Integralbegriff beschränken. Wir behandeln nur einfache Integrale von Funktionen einer reellen Variablen, wo also $f(x)$ in einer linearen Punktmenge A definiert ist und eine ebene Ordinatenmenge liefert; diese Einschränkung, die der bequemeren Ausdrucksweise dient, ist übrigens durchaus unerheblich und kann mit einem Federstrich beseitigt werden, indem man unter A eine n-dimensionale, unter B die entsprechende $(n+1)$-dimensionale euklidische Punktmenge versteht. Mehrfache Integrale (die bekanntlich von mehrmaligen einfachen Integralen zu unterscheiden sind) bieten also prinzipiell keine größere Schwierigkeit als einfache. Wesentlicher ist die im folgenden innegehaltene Einschränkung auf eigentliche Integrale, wo $f(x)$ eine beschränkte Funktion eines beschränkten Arguments ist und zu einer beschränkten Ordinatenmenge führt.

Wir bezeichnen die Längenmaße linearer Mengen mit $\varphi_1(A)$, $\psi_1(A)$, $f_1(A)$ (äußeres Maß, inneres Maß, Maß), die Flächenmaße ebener Mengen mit $\varphi_2(B)$, $\psi_2(B)$, $f_2(B)$.

Wir wählen in einer Ebene ein rechtwinkliges Achsensystem;

---

[1] Analytische Operationen sind, wenn man dem Ausdruck einen präzisen Sinn geben will, wohl die vier Spezies und die Bestimmung des Grenzwerts einer konvergenten Folge zu nennen; analytisch darstellbar diejenigen Funktionen, die aus x (oder mehreren Variablen) und Konstanten durch analytische Operationen entstehen. Alle diese Funktionen sind jedenfalls Bairesche Funktionen, also meßbar.

A sei eine beschränkte Punktmenge auf der Abszissenachse und $f(x)$ eine in A definierte, reelle, beschränkte, nichtnegative Funktion $(f(x) \geqq 0)$. Unter einer zu dieser Funktion gehörigen Ordinate B_x verstehen wir die Verbindungsstrecke der beiden Punkte $(x, 0)$ und $(x, f(x))$, wenn x ein Punkt von A ist; wir wollen dieser Strecke ihren oberen Endpunkt nicht zurechnen, also B_x als Menge der Punkte (x, y) mit festem x und

$$0 \leqq y < f(x)$$

erklären. B_x ist nur für positives $f(x)$ von Null verschieden. Die (beschränkte) Summe

$$B = \sum_x^A B_x$$

aller dieser Ordinaten nennen wir die zur Funktion gehörige Ordinatenmenge (der Leser kennt sie aus den Elementen als „das von der Kurve $y = f(x)$, der Abszissenachse und zwei Ordinaten begrenzte Flächenstück", wenn A nämlich ein Intervall ist). Wir definieren dann als oberes und unteres Integral[1] von $f(x)$ das äußere und innere Maß von B:

$$\overline{\int} f(x) = \varphi_2(B), \qquad \underline{\int} f(x) = \psi_2(B),$$

und als Integral schlechthin den gemeinsamen Wert beider im Gleichheitsfalle, also das Maß von B:

$$\int f(x) = f_2(B).$$

Wenn das Integral existiert, heißt die Funktion integrabel oder auch summierbar.[2]

Diese Erklärungen bezogen sich auf den Fall einer nichtnegativen Funktion. Ist $f(x)$ jetzt eine beliebige, aber beschränkte, in A definierte Funktion, so setzen wir

$$f'(x) = \tfrac{1}{2}\,|f(x)| + \tfrac{1}{2} f(x), \quad f''(x) = \tfrac{1}{2}\,|f(x)| - \tfrac{1}{2} f(x),$$

also

$$f(x) = f'(x) - f''(x).$$

Die Funktionen $f'(x)$, $f''(x)$ sind nichtnegativ; $f'(x)$ ist $= f(x)$, wo $f(x) \geqq 0$, sonst $= 0$, und Entsprechendes gilt von $f''(x)$. Wir

---

[1] Wir schreiben hier unbedenklich $\int f(x)$ statt $\int f(x)\,dx$ und unterdrücken auch die Angabe des Argumentes A (des „Integrationsgebietes").

[2] Dieser Ausdruck soll die Integrabilität im Lebesgueschen Sinne von der im Riemannschen Sinne unterscheiden.

definieren dann, wenn B', B'' die Ordinatenmengen dieser Funktionen sind, als oberes und unteres Integral von $f(x)$

$$\overline{\int} f(x) = \overline{\int} f'(x) - \underline{\int} f''(x) = \varphi_2(B') - \psi_2(B''),$$

$$\underline{\int} f(x) = \underline{\int} f'(x) - \overline{\int} f''(x) = \psi_2(B') - \varphi_2(B''),$$

woraus folgt

$$\overline{\int} f(x) - \underline{\int} f(x) = [\varphi_2(B') - \psi_2(B')] + [\varphi_2(B'') - \psi_2(B'')],$$

sodaß das obere Integral nicht kleiner als das untere ist. Beide fallen dann und nur dann zusammen, wenn einzeln B' und B'' meßbar, $f'(x)$ und $f''(x)$ integrabel sind, und der gemeinsame Wert wird in diesem Falle als Integral von $f(x)$

$$\int f(x) = \int f'(x) - \int f''(x) = f_2(B') - f_2(B'')$$

erklärt.

Der Funktion $|f(x)|$ entspricht die Ordinatenmenge $B = B' + B''$; zugleich mit $f(x)$ ist also $|f(x)|$ integrabel (nicht umgekehrt). Für die Funktion $-f(x)$ tauschen B' und B'' die Rollen, also ist

$$\overline{\int} - f(x) = - \underline{\int} f(x), \quad \underline{\int} - f(x) = - \overline{\int} f(x).$$

Ist $0 \leqq f(x) \leqq g(x)$, so ist die Ordinatenmenge von $f(x)$ in der von $g(x)$ enthalten, also

$$\overline{\int} f(x) \leqq \overline{\int} g(x), \quad \underline{\int} f(x) \leqq \underline{\int} g(x).$$

Diese Ungleichungen bleiben auch bestehen, wenn $f(x) \leqq g(x)$ von beliebigem Vorzeichen sind, da dann $f'(x) \leqq g'(x)$, $f''(x) \geqq g''(x)$. Insbesondere ist

$$\overline{\int} \pm f(x) \leqq \overline{\int} |f(x)|, \quad \underline{\int} \pm f(x) \leqq \underline{\int} |f(x)|,$$

woraus folgt, daß die Beträge von $\overline{\int} f(x)$ und $\underline{\int} f(x)$ höchstens gleich $\overline{\int} |f(x)|$ sind. Ist $f(x)$ integrabel, so ist

$$\left|\int f(x)\right| \leqq \int |f(x)|.$$

Werfen wir im Vorbeigehen noch einen Blick auf das Riemannsche Integral, das man erhält, wenn man nicht die Lebesgueschen Maße, sondern die Peano-Jordanschen Inhalte für φ_2, ψ_2, f_2 setzt; oberes und unteres Integral werden hier auch als die Darbouxschen Integrale bezeichnet. Es bedarf keiner Ausführung, daß durch den Übergang vom Inhalt zum Maß das obere Integral

(wenn es sich überhaupt ändert) verkleinert, das untere vergrößert, beide einander genähert werden. Das System der integrablen Funktionen erweitert sich wie das der meßbaren Mengen; eine Funktion mit Riemannschem Integral hat auch ein Lebesguesches (mit dem gleichen Wert), aber sie kann dieses ohne jenes haben. Dem Leser wird es vielleicht noch erwünscht sein, wenn wir die hier gegebene Erklärung der Darbouxschen Integrale mit der traditionellen zu völliger Übereinstimmung bringen. Nehmen wir $f(x) \geqq 0$ und A als abgeschlossenes Intervall (a, b) an (dies ist erlaubt, da man zu dem ursprünglichen Argument beliebig viele Stellen mit $f(x) = 0$ hinzufügen kann, die ja zur Ordinatenmenge keinen Beitrag liefern). Teilen wir das Intervall durch Einschaltung von Zwischenpunkten

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

und errichten über dem Teilintervall $x_{i-1} \leqq x \leqq x_i$ ein (abgeschlossenes) Rechteck R_i; die Summe dieser Rechtecke sei

$$R = \mathfrak{S}(R_1, R_2, \dots, R_n).$$

Durch ganz elementare Überlegungen erkennt man, daß $\varphi_2(B)$ untere Schranke der Inhalte $f_2(R)$ der Rechteckssummen $R \geqq B$, $\psi_2(B)$ obere Schranke der Inhalte $f_2(R)$ der Rechteckssummen $R \leqq B$ ist (d. h. daß man statt der definitionsgemäß zu verwendenden allgemeinen Polygone diese speziellen Polygone R benutzen darf). Wählt man noch die Höhen der Rechtecke im ersten Fall möglichst klein, im zweiten möglichst groß, so heißt das: ist μ_i die obere, λ_i die untere Schranke von $f(x)$ im Intervall (x_{i-1}, x_i) von der Länge $\delta_i = x_i - x_{i-1}$, so ist das obere Darbouxsche Integral $\overline{\int} f(x)$ die untere Schranke der Summen $\Sigma \mu_i \delta_i$, das untere $\underline{\int} f(x)$ die obere Schranke der Summen $\Sigma \lambda_i \delta_i$, für alle möglichen Intervallteilungen. Das ist die bekannte Definition in ihrer einfachsten Gestalt, und dies bleibt auch für Funktionen beliebigen Vorzeichens richtig.

Unsere weitere Betrachtung bezieht sich wieder auf das Lebesguesche Integral. Wir definieren wie in § 4, IV für eine reelle Zahl y

(1) $$A(y) = \text{Menge der Stellen } x, \text{ wo } f(x) \geqq y,$$

und nennen die Funktion $f(x)$ meßbar, wenn die Menge $A(y)$ für jedes y linear meßbar ist, d. h. ein Längenmaß $f_1(A(y))$ besitzt.

Integrale konstanter Funktionen. Ist h eine positive Zahl und $f(x) = h$ in A konstant, so ist

(2) $$\overline{\int} h = h \cdot \varphi_1(A), \qquad \underline{\int} h = h \cdot \psi_1(A).$$

Beweis. Ist zunächst G ein lineares Gebiet, H die zugehörige (der konstanten Funktion $f(x) = h$ entsprechende) Ordinatenmenge, so ist

$$f_2(H) = h \cdot f_1(G).$$

Denn G zerfällt in (endlich oder abzählbar viele) paarweise fremde, offene Strecken, H entsprechend in paarweise fremde Rechtecke von der Höhe h (mit einem Teil des Randes). Das Längenmaß von G ist die Summe der Streckenlängen, das Flächenmaß von H die Summe der Rechtecksflächen oder der mit h multiplizierten Streckenlängen.

Schließen wir jetzt A in ein lineares Gebiet G ein mit $f_1(G) < \varphi_1(A) + \varepsilon$, so schließt die Ordinatenmenge H von G die Ordinatenmenge B von A ein und es ist

$$\varphi_2(B) \leqq f_2(H) = h \cdot f_1(G) < h \cdot \varphi_1(A) + h\varepsilon,$$

also $\varphi_2(B) \leqq h \cdot \varphi_1(A)$.

Andererseits sei B' die Menge, die aus B durch Hinzufügung der oberen Endpunkte der Ordinaten entsteht $(0 \leqq y \leqq h)$; B' unterscheidet sich von B um eine Menge vom Flächenmaß Null und es ist daher $\varphi_2(B') = \varphi_2(B)$. Schließen wir B' in ein ebenes Gebiet H' ein mit $f_2(H') < \varphi_2(B') + \varepsilon$. Zu jedem Punkt z der Ordinate B'_x gibt es eine quadratische Umgebung $Q_z \leqq H'$ (ein Quadratinneres mit z als Mittelpunkt und Seiten parallel den Koordinatenachsen); nach dem Borelschen Satze genügen, da B'_x abgeschlossen ist, endlich viele Q_z zur Bedeckung von B'_x, und aus diesen Q_z gewinnt man ein in H' liegendes, B'_x einschließendes unberandetes Rechteck H'_x mit Seiten parallel den Koordinatenachsen. Von diesem Rechteck, das unter die Abszissenachse und über die Gerade $y = h$ hinausragt, sei G_x der Durchschnitt mit der Abszissenachse und H_x der Durchschnitt mit dem Parallelstreifen $0 \leqq y < h$. Dann ist G_x eine offene Strecke, $G = \mathfrak{S}\, G_x$ ein lineares Gebiet $\geqq A$ und $H = \mathfrak{S}\, H_x \leqq H'$ die Ordinatenmenge zu G. Demnach hat man

$$\varphi_2(B) + \varepsilon > f_2(H') \geqq f_2(H) = h \cdot f_1(G) \geqq h \cdot \varphi_1(A),$$

also $\varphi_2(B) \geqq h \cdot \varphi_1(A)$ und schließlich $\varphi_2(B) = h \cdot \varphi_1(A)$.

Durch Komplementbildung (indem man A in eine Strecke, B in das entsprechende Rechteck einschließt) findet man $\psi_2(B) = h \cdot \psi_1(A)$, womit die Gleichungen (2) bewiesen sind. Das Integral $\int h$ existiert dann und nur dann, wenn A meßbar ist, und ist $= h \cdot f_1(A)$; dies gilt auch für negatives h.

Zerlegung in Horizontalstreifen. Für eine in A definierte nichtnegative Funktion $f(x)$ schneiden wir die Ordinatenmenge $B = \Sigma B_x$ mit dem Parallelstreifen $h \leqq y < k$, wobei $h \geqq 0$ voraus-

gesetzt sei; der Durchschnitt sei $D = \Sigma D_x$. Der Ordinatendurchschnitt D_x ist durch das Zusammenbestehen der Ungleichungen

$$h \leqq y < k, \; f(x)$$

definiert, verschwindet also jedenfalls für $f(x) < h$ (sogar auch noch für $f(x) = h$); und D kann, wenn wir die Abszissenachse nach $y = h$ verlegen, als Ordinatenmenge einer Funktion angesehen werden, die nur für $f(x) \geqq h$, also nach (1) in der Menge $A(h)$ definiert, in der Menge $A(k)$ gleich $k - h$ und sonst gleich

$$f(x) - h < k - h$$

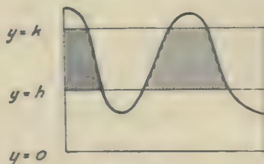


Fig. 53.

ist. D liegt also zwischen zwei Mengen, die sich als Ordinatenmengen der konstanten Funktion $k - h$ mit den Argumenten $A(h)$ und $A(k)$ auffassen lassen, und daraus folgt

$$(3) \quad \begin{cases} (k-h)\,\varphi_1\big(A(k)\big) \leqq \varphi_2(D) \leqq (k-h)\,\varphi_1\big(A(h)\big), \\ (k-h)\,\psi_1\big(A(k)\big) \leqq \psi_2(D) \leqq (k-h)\,\psi_1\big(A(h)\big). \end{cases}$$

Aus diesen Formeln gewinnen wir eine notwendige Integrabilitätsbedingung. Ist $f(x)$ integrabel, so ist B, also auch D meßbar, $\varphi_2(D) = \psi_2(D)$, und die letzten Ungleichungen geben

$$\varphi_1\big(A(k)\big) \leqq \psi_1\big(A(h)\big) \qquad (0 \leqq h < k).$$

Halten wir k fest und lassen h eine Folge positiver, wachsender, nach k konvergenter Zahlen h_n durchlaufen, so bilden die Mengen $A(h_n)$ eine absteigende Folge mit dem Durchschnitt $A(k)$ und es ist (§ 3, (5))

$$\psi_1\big(A(k)\big) = \lim \psi_1\big(A(h_n)\big) \geqq \varphi_1\big(A(k)\big),$$

also $\psi_1\big(A(k)\big) = \varphi_1\big(A(k)\big)$, mit andern Worten:

Damit $f(x)$ integrabel sei, muß jede Menge $A(y)$ für $y > 0$ linear meßbar sein.

Für die Menge $A(0) = A$ versagt dieser Schluß, und das ist ganz naturgemäß, da ja die Stellen, wo $f(x) = 0$, keinen Beitrag zur Ordinatenmenge geben; aus eben demselben Grunde kann man allerdings (nach eventueller Hinzufügung solcher Stellen) A als meßbar, z. B. als Intervall, voraussetzen. Da auch, für negatives y, $A(y) = A$ ist, so ergibt sich: eine (in einer meßbaren Menge A definierte) integrable Funktion ist meßbar.

Wenn das Integral der nichtnegativen Funktion $f(x)$ Null ist, so ist $f(x) = 0$ bis auf eine Menge vom Maße Null. Denn dann ist $\varphi_2(B) = 0$, also in (3) $\varphi_2(D) = 0$, $\varphi_1\big(A(k)\big) = 0$; die Menge $A(k)$ für positives k und folglich die Menge

$$\mathfrak{S}\big(A(1),\; A(\tfrac{1}{2}),\; A(\tfrac{1}{3}),\; \ldots\big)$$

der Stellen, wo $f(x) > 0$, hat das Längenmaß Null.

Um zu beweisen, daß auch umgekehrt eine meßbare Funktion integrabel ist, verstehen wir unter d die obere Schranke von $f(x)$ oder eine noch größere Zahl, sodaß die Ordinatenmenge im Horizontalstreifen $0 \leqq y < d$ liegt. Wir teilen durch Einschaltung von Zahlen

$$0 = y_0 < y_1 < y_2 < \cdots < y_{n-1} < y_n = d$$

diesen Streifen in Teilstreifen, und es sei B_i der Durchschnitt von B mit dem Streifen

$$y_{i-1} \leqq y < y_i \qquad (i = 1, 2, \ldots, n).$$

Hierdurch wird $B = \Sigma B_i$ in relativ meßbare Mengen zerlegt, sodaß (S. 415)

$$\varphi_2(B) = \Sigma \varphi_2(B_i), \quad \psi_2(B) = \Sigma \psi_2(B_i).$$

Benutzt man dann die Formeln (3), so folgt

$$\Sigma(y_i - y_{i-1})\varphi_1\big(A(y_i)\big) \leqq \varphi_2(B) \leqq \Sigma(y_i - y_{i-1})\varphi_1\big(A(y_{i-1})\big)$$

und analog für ψ.

Da $\varphi_1\big(A(y)\big)$ eine mit wachsendem y monoton abnehmende (genauer: nicht zunehmende) Funktion von y, also im Riemannschen Sinne integrabel ist, so ist, bei allen möglichen Zerlegungen in Teilstreifen, die obere Schranke der linken Seite und die untere der rechten Seite in der letzten Formel das Riemannsche Integral von $\varphi_1\big(A(y)\big)$, also in gewöhnlicher Schreibweise

$$\text{(4)} \qquad \varphi_2(B) = \int_0^d \varphi_1\big(A(y)\big)\,dy, \quad \psi_2(B) = \int_0^d \psi_1\big(A(y)\big)\,dy.$$

Hiermit sind Lebesguesche Integrale auf gewöhnliche, Flächenmaße auf Längenmaße zurückgeführt; zugleich ergibt sich, daß eine meßbare Funktion integrabel und

$$\text{(5)} \qquad f_2(B) = \int_0^d f_1\big(A(y)\big)\,dy$$

ihr Integral ist. Wenn $f(x)$ bis auf eine Menge vom Längenmaß Null verschwindet, also $f_1\big(A(y)\big) = 0$ für positives y, so ist $\int f(x) = 0$.

Falls es sich, statt um eine nichtnegative, um eine (beschränkte) Funktion $f(x)$ beliebigen Vorzeichens handelt und wir das Argument A als meßbar annehmen, so lehrt die Definition des Integrals (S. 432): ist $f(x)$ integrabel, so sind $f'(x)$ und $f''(x)$ integrabel, also meßbar, und $f(x) = f'(x) - f''(x)$ ebenfalls meßbar; ist $f(x)$ meßbar, so sind auch $|f(x)|$, $f'(x)$ und $f''(x)$ meßbar, also $f'(x)$ und $f''(x)$ und damit $f(x)$ integrabel. Die Identität integrabler und meßbarer (beschränkter) Funktionen gilt also für Funktionen beliebigen Vorzeichens.

Zerlegung in Vertikalstreifen. Eine in A definierte Funktion $f(x)$ gibt für jede Teilmenge P von A eine Teilfunktion $f(x \mid P)$, deren Integrale (über das Argument P) wir mit

$$\overline{\int} f(x \mid P), \quad \underline{\int} f(x \mid P), \quad \int f(x \mid P)$$

bezeichnen. Diese Integrale sind für nichtnegatives $f(x)$

$$\varphi_2(Q), \ \psi_2(Q), \ f_2(Q),$$

wenn $Q = \sum\limits_x^P B_x$ die der Teilfunktion entsprechende Ordinatenmenge ist. Wenn P in A meßbar ist, so ist auch Q in B meßbar; denn ist $P = \mathfrak{D}(A, M)$, so ist $Q = \mathfrak{D}(B, N)$, wo N eine über der linear meßbaren Menge M errichtete Ordinatenmenge von konstanter, hinreichend großer Höhe ist. Wird A in eine endliche oder abzählbare Menge paarweise fremder Summanden $A = A_1 + A_2 + \ldots$ zerlegt, so entspricht dieser eine Zerlegung $B = B_1 + B_2 + \ldots$ der Ordinatenmenge, und aus den Formeln § 3, (6), (7) folgt

$$\overline{\int} f(x) \leqq \Sigma \overline{\int} f(x \mid A_n), \quad \underline{\int} f(x) \geqq \Sigma \underline{\int} f(x \mid A_n).$$

Nehmen wir aber insbesondere die A_n in A meßbar an, so sind auch die B_n in B meßbar und es gelten die Gleichungen

$$\overline{\int} f(x) = \Sigma \overline{\int} f(x \mid A_n), \quad \underline{\int} f(x) = \Sigma \underline{\int} f(x \mid A_n), \tag{6}$$

von denen man unmittelbar sieht, daß sie, und zwar mit absoluter Konvergenz der Reihen rechterhand, für eine Funktion beliebigen Vorzeichens bestehen bleiben. Ist $f(x)$ integrabel, so auch jedes $f(x \mid A_n)$, und dann gilt

$$\int f(x) = \Sigma \int f(x \mid A_n). \tag{7}$$

Insbesondere ist dies alles richtig, wenn A und die A_n (absolut) meßbar sind.

Um hiervon eine Anwendung zu machen, nehmen wir A als Intervall (a, b) und $f(x)$ als integrabel an; unter $\int\limits_\alpha^\beta f(x)$ verstehen wir das Integral über ein Teilintervall (α, β) für $\alpha \leqq \beta$ und setzen

$$\int\limits_\beta^\alpha f(x) = -\int\limits_\alpha^\beta f(x).$$

Ob wir die Endpunkte des Intervalls mitzählen, ist natürlich belanglos, da einzelne Ordinaten das Flächenmaß 0 haben. Sind alle diese Intervallintegrale (oder nur die bei fester unterer und variabler oberer Grenze) bekannt, so ist $\int f(x \mid P)$ für jede meßbare

Teilmenge P von A bekannt; denn nach (7) erhält man das Integral über ein Gebiet G als Summe von Integralen über paarweise fremde Intervalle, das Integral über eine abgeschlossene Menge F als Differenz von Integralen über A und G ($A - F$ ist, von den Endpunkten von A eventuell abgesehen, ein G), das Integral über eine Menge F_σ wieder nach (7), da wir mit aufsteigenden F_n

$$F_\sigma = \mathfrak{S}(F_1, F_2, \ldots) = F_1 + (F_2 - F_1) + (F_3 - F_2) + \ldots$$

annehmen können, endlich für eine beliebige meßbare Menge P, die wir ja in ein F_σ mit gleichem Maß und eine Menge vom Maße Null zerlegen können, $\int f(x|P) = \int f(x|F_\sigma)$. Wenn insbesondere alle Intervallintegrale verschwinden, so auch alle Integrale über meßbare Teilmengen; dann muß aber $f(x)$ bis auf eine Menge vom Maße Null verschwinden. Denn ist P die meßbare Menge der Stellen, wo $f(x) > 0$, so folgt aus dem Verschwinden von $\int f(x|P)$ nach S. 435, daß P vom Maße Null sein muß, und das Gleiche für die Menge der Stellen mit $f(x) < 0$. Also: eine in (a, b) integrable Funktion $f(x)$, deren Integral in jedem Teilintervall verschwindet, verschwindet selbst bis auf eine Menge vom Maße Null.

Kehren wir zu den Formeln (6) zurück; wir wollen jetzt A als meßbar und eine Zerlegung in endlich viele meßbare Mengen

$$A = A_1 + A_2 + \cdots + A_n = \Sigma A_i$$

annehmen; zur Abkürzung setzen wir

$$\delta_i = f_1(A_i), \quad \Sigma \delta_i = f_1(A).$$

Ist weiter μ_i die obere, λ_i die untere Schranke von $f(x)$ in A_i, so gibt die Ungleichung $\lambda_i \leqq f(x \mid A_i) \leqq \mu_i$ durch Integration über A_i

$$\lambda_i \delta_i \leqq \underline{\int} f(x \mid A_i) \leqq \overline{\int} f(x \mid A_i) \leqq \mu_i \delta_i,$$

(8) $$\Sigma \lambda_i \delta_i \leqq \underline{\int} f(x) \leqq \overline{\int} f(x) \leqq \Sigma \mu_i \delta_i.$$

Nennt man Φ die untere Schranke der oberen Summen $\Sigma \mu_i \delta_i$, Ψ die obere Schranke der unteren Summen $\Sigma \lambda_i \delta_i$ bei allen möglichen Zerlegungen, so ist also

(9) $$\overline{\int} f(x) \leqq \Phi, \quad \underline{\int} f(x) \geqq \Psi.$$

Wir werden nachher sehen, daß hierin die Gleichheitszeichen gelten, sodaß wir damit die beiden Lebesgueschen Integrale in formal ähnlicher Weise wie die Darbouxschen erklären können, nur daß statt der speziellen Zerlegungen in Intervalle die allgemeinen in meßbare Mengen zu betrachten sind.

Ist zunächst $f(x)$ meßbar oder integrabel, so gehört zu den fraglichen Zerlegungen die folgende: wenn alle Werte von $f(x)$ zwischen c und d liegen, so teilen wir dies Intervall wieder durch Einschaltung von Zwischenpunkten

$$c = y_0 < y_1 < y_2 < \dots < y_{n-1} < y_n = d$$

und setzen

$$A_i = A(y_{i-1}) - A(y_i),$$

d. h.

$$A_i = \text{Menge der Stellen, wo } y_{i-1} \leqq f(x) < y_i;$$

denn diese Mengen sind meßbar. Da dann $\mu_i \leqq y_i$, $\lambda_i \geqq y_{i-1}$, so kann die Differenz der oberen und unteren Summe

$$\Sigma(\mu_i - \lambda_i)\delta_i \leqq \Sigma(y_i - y_{i-1})\delta_i$$

beliebig klein gemacht werden, indem man die Teilintervalle $y_i - y_{i-1} < \varepsilon$ wählt (die Summe wird dann $< \varepsilon \cdot \Sigma\delta_i = \varepsilon \cdot f_1(A)$). Für eine integrable Funktion ist also in der Tat

$$\int f(x) = \Phi = \Psi,$$

das Integral die untere Schranke der oberen und zugleich die obere Schranke der unteren Summen.

Sind $f(x)$ und $g(x)$, also auch $h(x) = f(x) + g(x)$ integrable Funktionen in A, so gilt für die bezüglichen Schranken in irgend einer Teilmenge von A

(10) $$\mu_h \leqq \mu_f + \mu_g, \quad \lambda_h \geqq \lambda_f + \lambda_g$$

und demnach

$$\int h(x) \leqq \int f(x) + \int g(x), \quad \int h(x) \geqq \int f(x) + \int g(x),$$

also

$$\int (f(x) + g(x)) = \int f(x) + \int g(x),$$

das Integral der Summe ist die Summe der Integrale.

Für eine Konstante c ist

$$\int cf(x) = c \cdot \int f(x).$$

Wenn die in A definierten, integrablen Funktionen $f_n(x)$ gleichmäßig beschränkt sind, d. h. eine Ungleichung

$$|f_n(x)| \leqq \mu$$

für alle n und x besteht, und wenn $f(x) = \lim f_n(x)$, so ist

(11) $$\int f(x) = \lim \int f_n(x),$$

das Integral des Limes gleich dem Limes des Integrals.

Es ist nämlich auch $f(x)$ integrabel (meßbar) und mit $g_n(x) = |f_n(x) - f(x)|$

$$\left|\int f_n(x) - \int f(x)\right| = \left|\int (f_n(x) - f(x))\right| \leqq \int g_n(x).$$

Für ein bestimmtes $\varepsilon > 0$ sei P_n die Menge der Stellen x, wo $g_n(x) < \varepsilon$, $Q_n = A - P_n$ die Menge derer, wo $g_n(x) \geqq \varepsilon$; in Q_n ist aber wenigstens $g_n(x) \leqq |f_n(x)| + |f(x)| \leqq 2\mu$. Da P_n und Q_n meßbar sind, so folgt

$$\int g_n(x) \leqq \varepsilon \cdot f_1(P_n) + 2\mu \cdot f_1(Q_n).$$

Nun ist $\operatorname{Lim\,sup} Q_n = 0$ (S. 427), $\lim f_1(Q_n) = 0$, also für hinlänglich großes n das zweite Glied der rechten Seite $< \varepsilon$ und

$$\int g_n(x) < \varepsilon \cdot f_1(A) + \varepsilon, \quad \lim \int g_n(x) = 0,$$

womit die Formel (11) bewiesen ist.

Die Tatsache, daß eine konvergente Folge integrabler Funktionen immer wieder eine integrable Funktion liefert, und die relativ geringe Einschränkung, unter der die Vertauschung von Integral und Limes (im Fall einer Funktionensumme also die Integration gliedweise) gestattet ist, bildet einen fundamentalen Vorzug der Lebesgueschen Integrale vor den Riemannschen, bei denen die gleichmäßige Konvergenz — eine freilich auch zu enge Bedingung — gefordert zu werden pflegt.

Um Anwendungen davon zu geben, sei B eine beschränkte ebene Punktmenge, nicht mehr notwendig eine Ordinatenmenge, B_x ihr Durchschnitt mit der zur Abszisse x gehörigen Geraden parallel der Ordinatenachse, A die Projektion von B auf die Abszissenachse oder eine diese Projektion enthaltende beschränkte Teilmenge der Abszissenachse, etwa ein hinreichend großes Intervall, so daß außerhalb A jedenfalls $B_x = 0$ und $B = \sum_x^A B_x$ ist. Wir suchen einen Zusammenhang zwischen dem Flächenmaß von B und den Längenmaßen der B_x.

Ist B zunächst ein Polygon P, so ist P_x eine Summe endlich vieler Strecken und

$$f_2(P) = \int f_1(P_x).$$

Denn die Ordinatenmenge zur Funktion $f_1(P_x)$ ist offenbar selbst ein Polygon, dessen Flächengleichheit mit P elementargeometrisch bewiesen werden kann, durch Zerlegung in rechtwinklige Dreiecke mit Katheten parallel den Koordinatenachsen.

Ist ferner G ein ebenes Gebiet, so läßt sich dies als Summe einer aufsteigenden Folge $P_1 \leqq P_2 \leqq \ldots$ von Polygonen darstellen; entsprechend ist das lineare Gebiet G_x die Summe der Folge

$P_{1x} \leqq P_{2x} \leqq \ldots$. Nach der Limeseigenschaft der Lebesgueschen Integrale ist

$$f_2(G) = \lim f_2(P_n) = \lim \int f_1(P_{nx})$$
$$= \int \lim f_1(P_{nx}) = \int f_1(G_x).$$

Entsprechendes gilt für eine abgeschlossene Menge F, die ja als Durchschnitt einer absteigenden Polygonfolge darstellbar ist. Wir haben also

$$f_2(G) = \int f_1(G_x), \quad f_2(F) = \int f_1(F_x).$$

Für eine beliebige Menge B bestimme man ein ebenes Gebiet $G \geqq B$ mit $f_2(G) < \varphi_2(B) + \varepsilon$; $G_x \geqq B_x$ ist ein lineares Gebiet, also $f_1(G_x) \geqq \varphi_1(B_x)$. Das gibt

$$\varphi_2(B) + \varepsilon > f_2(G) = \int f_1(G_x) = \overline{\int} f_1(G_x) \geqq \overline{\int} \varphi_1(B_x)$$

und die erste der Ungleichungen

(12) $$\varphi_2(B) \geqq \overline{\int} \varphi_1(B_x), \qquad \psi_2(B) \leqq \underline{\int} \psi_1(B_x);$$

die zweite erhält man durch entsprechende Verwendung einer Menge $F \leqq B$. Ist B meßbar, so folgt

$$\overline{\int} \varphi_1(B_x) \leqq \underline{\int} \psi_1(B_x),$$

also die Gleichheit dieser beiden Integrale und der beiden andern; d. h. $\varphi_1(B_x)$ und $\psi_1(B_x)$ sind integrabel und

(13) $$f_2(B) = \int \varphi_1(B_x) = \int \psi_1(B_x).$$

Approximation durch integrable Funktionen. Zu jeder Funktion $f(x)$ gibt es eine integrable Funktion $\bar{f}(x) \geqq f(x)$ und eine integrable Funktion $\underline{f}(x) \leqq f(x)$ mit

(14) $$\overline{\int} f(x) = \int \bar{f}(x), \quad \underline{\int} f(x) = \int \underline{f}(x).$$

Denn ist $f(x)$ in A definiert und zunächst $\geqq 0$, so bestimme man zu seiner Ordinatenmenge B eine meßbare Menge $M \geqq B$ (z. B. ein G_δ) mit $f_2(M) = \varphi_2(B)$. Man kann dieses M durch jede meßbare Menge zwischen M und B ersetzen, so durch die (bei linear meßbarem A eben meßbare) Menge $\sum\limits_x^A M_x$; die Funktion

$$\varphi_1(M_x) \geqq \varphi_1(B_x) = f(x)$$

ist integrabel und ihr Integral $= \varphi_2(B)$. Die erste und analog die zweite Behauptung gilt also für eine nichtnegative Funktion $f(x)$; man überträgt sie unmittelbar auf eine Funktion beliebigen Vorzeichens.

Damit ist unmittelbar zu erkennen, daß in den Formeln (9) die Gleichheitszeichen gelten, also $\overline{\int} f$ die untere Schranke Φ der oberen Summen $\Sigma\, \mu_i \delta_i$ und $\underline{\int} f$ die obere Schranke Ψ der unteren Summen $\Sigma\, \lambda_i \delta_i$ ist. Denn $\overline{\int} f = \int \bar{f}$ ist die untere Schranke der oberen Summen $\Sigma\, \bar{\mu}_i \delta_i \geqq \mu_i \delta_i$, also $\geqq \Phi$, was mit (9) kombiniert $\overline{\int} f = \Phi$ ergibt; ebenso $\underline{\int} f = \Psi$.

Hieraus mit Hilfe von (10) oder durch Verwendung approximierender integrabler Funktionen findet man die Ungleichungen

$$(15) \qquad \left\{ \begin{aligned} \overline{\int}(f+g) &\leqq \overline{\int} f + \overline{\int} g, \\ \underline{\int}(f+g) &\geqq \underline{\int} f + \underline{\int} g. \end{aligned} \right.$$

Vertauscht man g mit $-g$, so liefert die erste

$$\overline{\int}(f-g) \leqq \overline{\int} f - \underline{\int} g$$

oder mit Ersetzung von f durch $f+g$ die erste der beiden Formeln

$$(16) \qquad \left\{ \begin{aligned} \overline{\int}(f+g) &\geqq \overline{\int} f + \underline{\int} g, \\ \underline{\int}(f+g) &\leqq \overline{\int} f + \underline{\int} g. \end{aligned} \right.$$

Diese Summenformeln für Funktionen entsprechen denen für Mengen (§ 3, (3)), was übrigens cum grano salis zu verstehen ist: sind B_1 und B_2 die Ordinatenmengen der nichtnegativen Funktionen $f_1(x)$ und $f_2(x)$, so ist $B = \mathfrak{S}(B_1, B_2)$ die Ordinatenmenge nicht von $f_1(x) + f_2(x)$, sondern von $f(x) = \max\,[f_1(x), f_2(x)]$, und $B' = \mathfrak{D}(B_1, B_2)$ die von $f'(x) = \min\,[f_1(x), f_2(x)]$.

Auf die Theorie der uneigentlichen Integrale, wo Argument und Funktion unbeschränkt sein können,[1] wollen wir nicht eingehen; wir bemerken nur, daß bei Zugrundelegung derselben Definitionen wie S. 432 mit $\int f(x)$ auch $\int |f(x)|$ existiert, also nur die sogenannten absolut konvergenten Integrale auf diese Weise in Betracht gezogen werden. Die weitere Behandlung ist der hier gegebenen analog, wenn auch nicht alles buchstäblich zu übertragen ist; z. B. wird man bei der Zerlegung in Horizontal- und Vertikalstreifen auch abzählbar viele Summanden berücksichtigen müssen.

---

[1] Man wird hier sogar zweckmäßig die uneigentlichen Funktionswerte $+\infty$, $-\infty$ zulassen; dem Wert $+\infty$ entspricht als Ordinate eine Halbgerade $y \geqq 0$.

§ 6. Differentiation und Integration.

In dem abgeschlossenen Intervall (a, b) sei eine reelle Funktion $f(x)$ definiert.[1] Man teile das Intervall durch Einschiebung von Zwischenpunkten

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

und bilde die Summe der Beträge der Funktionsdifferenzen

$$v_0 = \sum |f(x_i) - f(x_{i-1})| \qquad (i = 1, 2, \dots, n).$$

Wenn diese v_0, für alle Intervallteilungen, eine beschränkte Menge bilden, so heißt die Funktion von beschränkter Variation; die obere Schranke v der v_0 heißt die (totale) Variation von $f(x)$ im Intervall (a, b).

Bezeichnet man mit p_0 die Summe der positiven, mit $-n_0$ die Summe der negativen unter den Differenzen $f(x_i) - f(x_{i-1})$ (natürlich $p_0 = 0$, wenn keine positiven existieren), so ist

$$v_0 = p_0 + n_0, \quad f(b) - f(a) = p_0 - n_0.$$

Die oberen Schranken p, n der Zahlen p_0, n_0 bei allen Intervallteilungen heißen die positive und negative Variation von $f(x)$. Da $f(b) - f(a)$ von der Intervallteilung unabhängig ist, so folgt aus

$$p_0 = \tfrac{1}{2} v_0 + \tfrac{1}{2}[f(b) - f(a)], \quad n_0 = \tfrac{1}{2} v_0 - \tfrac{1}{2}[f(b) - f(a)],$$

daß diese Gleichungen auch beim Übergang zu den oberen Schranken bestehen bleiben, also

$$v = p + n, \quad f(b) - f(a) = p - n.$$

Ist $f(x)$ im Intervall (a, b) von beschränkter Variation, so auch in jedem Intervall (a, x) für $a \leqq x \leqq b$. Werden die totale, positive, negative Variation in diesem Intervall mit $v(x)$, $p(x)$, $n(x)$ bezeichnet, so sind dies offenbar monoton wachsende Funktionen, d. h. $v(x) \leqq v(x')$ für $x < x'$. Es folgt daraus, daß eine Funktion beschränkter Variation

$$f(x) = f(a) + p(x) - n(x)$$

als Differenz (oder Summe) zweier monotoner Funktionen dargestellt werden kann, übrigens, wie man unmittelbar erkennt, eine stetige Funktion beschränkter Variation als Differenz stetiger monotoner Funktionen. Umgekehrt ist eine monotone Funktion jedenfalls von beschränkter Variation, da hier stets $v_0 = |f(b) - f(a)|$ ist, und die Summe oder Differenz zweier Funktionen von beschränkter Variation ist wieder eine solche.

---

[1] In diesem Paragraphen ist die Beschränkung auf Funktionen einer reellen Variablen nicht nur formaler Natur.

Bilden wir eine wohlgeordnete, abgeschlossene Menge

$$X = \{x_0, x_1, \ldots, x_\xi, \ldots, x_\eta\}$$

mit $x_0 = a$, $x_\eta = b$. Ist $f(x)$ in (a, b) stetig und von beschränkter Variation v, so ist

(1) $$f(b) - f(a) = \sum_\xi [f(x_{\xi+1}) - f(x_\xi)] \qquad (\xi < \eta).$$

Für endliches η ist dies evident. Ist η unendlich, so ist es abzählbar, da jede wohlgeordnete lineare Punktmenge höchstens abzählbar ist (S. 279). Da nun die Summe der Beträge endlich vieler Glieder der Reihe rechterhand stets $\leqq v$ ist, so ist diese Reihe absolut konvergent, ebenso ihre Teilreihen. Hiernach ist durch Induktion leicht zu beweisen, daß die behauptete Formel für die angenommene Menge X vom Typus $\eta + 1$ gilt, wenn sie für alle Mengen des Typus $\xi + 1$ ($\xi < \eta$) richtig ist. Der Schluß von η auf $\eta + 1$ ist evident. Ist η hingegen Limeszahl und zwar Limes der Folge $0 = \eta_0 < \eta_1 < \eta_2 < \ldots$, so erhält man durch Spaltung in Teilreihen (die wegen der absoluten Konvergenz zulässig ist)

$$\sum_\xi = \sum_{\xi_1} + \sum_{\xi_2} + \cdots,$$

wobei diese Indices gemäß den Ungleichungen

$$\eta_0 \leqq \xi_1 < \eta_1, \ \eta_1 \leqq \xi_2 < \eta_2, \ \ldots$$

variieren, also

$$\begin{aligned} \sum_\xi &= [f(x_{\eta_1}) - f(x_{\eta_0})] + [f(x_{\eta_2}) - f(x_{\eta_1})] + \cdots \\ &= \lim f(x_{\eta_n}) - f(x_0). \end{aligned}$$

Da nun die Menge X abgeschlossen war, so ist $\lim x_{\eta_n} = x_\eta$, und da $f(x)$ stetig sein sollte, $\lim f(x_{\eta_n}) = f(x_\eta)$, mithin

$$\sum_\xi = f(x_\eta) - f(x_0) = f(b) - f(a).$$

Sei jetzt $f(x)$ eine im Intervall $A = (a, b)$ definierte Funktion, deren sämtliche Differenzenquotienten

$$f(\alpha, \beta) = \frac{f(\beta) - f(\alpha)}{\beta - \alpha}$$

eine beschränkte Menge bilden; $f(x)$ ist offenbar stetig und von beschränkter Variation. Ihre vier Ableitungen (S. 396) sind beschränkt und meßbar, also integrabel. Sei etwa

$$g(x) = \lim_{h=0} \sup f(x, x+h) \qquad (h > 0)$$

die obere rechte Ableitung ($g(b)$ ein beliebiger Wert); wir wollen zeigen, daß

(2) $$\int g(x) = f(b) - f(a).$$

Nach dem Begriff des oberen Limes gibt es für jedes x außer b und jedes positive ε beliebig kleine (positive) Werte von h, für die

$$f(x, x+h) > g(x) - \varepsilon,$$

während zugleich für alle hinreichend kleinen Werte von h

$$f(x, x+h) < g(x) + \varepsilon;$$

es gibt also beliebig kleine h, die beide Ungleichungen erfüllen.

Wählen wir die positive Zahl μ so groß, daß für alle x

$$|g(x)| < \mu$$

und teilen das Intervall $(-\mu, \mu)$ in n der Einfachheit wegen gleichlange Teilintervalle (y_{i-1}, y_i) mit $y_i = -\mu + \frac{i}{n} \cdot 2\mu$; ist A_i die Menge der Stellen, wo

$$y_{i-1} < g(x) \leqq y_i, \qquad (i = 1, 2, \ldots, n)$$

so ist

$$(3) \qquad \sum y_{i-1} f_1(A_i) \leqq \int g(x) \leqq \sum y_i f_1(A_i).$$

Die äußeren Glieder unterscheiden sich wegen

$$A = \sum A_i, \quad f_1(A) = \sum f_1(A_i) = b - a$$

um $\frac{2\mu}{n}(b-a)$, also um beliebig wenig für hinreichend großes n. Wir schließen ferner jede Menge A_i in ein lineares Gebiet G_i ein mit

$$f_1(G_i) < f_1(A_i) + \frac{1}{n^2}.$$

Ist $a \leqq x < b$, und gehört x zu A_i, so gibt es wegen $y_{i-1} < g(x) < y_{i+1}$ beliebig kleine positive h derart, daß $x + h \leqq b$ und

$$y_{i-1} < f(x, x+h) < y_{i+1};$$

wir wählen unter diesen ein (von x abhängiges) $h = h(x)$ derart, daß das Intervall

$$J(x) = \text{Menge der Punkte } u \text{ mit } x \leqq u < x + h(x)$$

noch dem Gebiet G_i angehört. Den rechten Endpunkt dieses dem Punkte x zugeordneten Intervalls bezeichnen wir noch mit

$$\varphi(x) = x + h(x).$$

Man kann nun mit einer endlichen oder abzählbaren Menge paarweise fremder Intervalle $J(x)$ das Gesamtintervall in folgender Weise bedecken. Wir weisen gewissen Ordnungszahlen $< \omega_1$ Punkte des Intervalls A zu, indem wir durch Induktion definieren: $x_0 = a$ und für $\eta > 0$

$$x_\eta = \text{obere Schranke von } \varphi(x_\xi) \text{ für } \xi < \eta,$$

vorausgesetzt, daß alle x_ξ definiert und $< b$ sind. Betrachtungen,

wie sie dem Leser nun schon geläufig sind, lehren, daß hiermit eine wohlgeordnete abgeschlossene Menge

$$X = \{x_0, x_1, \ldots, x_\xi, \ldots, x_\eta\}$$

mit $x_0 = a$ und $x_\eta = b$ definiert wird; ihr Anfang lautet, falls sie unendlich ist,

$$x_0 = a,\ x_1 = \varphi(x_0) > x_0,\ x_2 = \varphi(x_1) > x_1,\ \ldots,\ x_\omega = \lim x_\nu,\ \ldots.$$

Für $\xi < \eta$ ist $J(x_\xi)$ das wie oben bestimmte Intervall zwischen x_ξ incl. und $\varphi(x_\xi) = x_{\xi+1}$ excl. Bedeutet noch J_i die Summe derjenigen Intervalle $J(x_\xi)$, deren linke Endpunkte x_ξ zu A_i gehören, so ist

$$A - \{b\} = \sum_\xi J(x_\xi) = \sum_i J_i;$$

jedes J_i ist als Intervallsumme meßbar und

$$\Sigma f_1(J_i) = b - a.$$

Nach Konstruktion ist nun, wenn x_ξ zu A_i gehört,

$$y_{i-1} < f(x_\xi, x_{\xi+1}) < y_{i+1},$$

$$y_{i-1}(x_{\xi+1} - x_\xi) < f(x_{\xi+1}) - f(x_\xi) < y_{i+1}(x_{\xi+1} - x_\xi).$$

Also durch Summation über alle $\xi < \eta$, wobei drei absolut konvergente Reihen entstehen, und auf Grund von (1)

$$\Sigma y_{i-1} f_1(J_i) \leqq f(b) - f(a) \leqq \Sigma y_{i+1} f_1(J_i); \tag{4}$$

die äußeren Glieder dieser Ungleichung unterscheiden sich um $\frac{4\mu}{n} \cdot (b - a)$.

Endlich vergleichen wir noch die linken Glieder in (3) und (4) mit einander und mit $\Sigma y_{i-1} f_1(G_i)$, wobei zu beachten ist, daß J_i und A_i Teilmengen von G_i sind. Demnach ist

$$|\Sigma y_{i-1}[f_1(G_i) - f_1(A_i)]| \leqq \mu \cdot \Sigma [f_1(G_i) - f_1(A_i)] < \frac{\mu}{n},$$

ebenso

$$|\Sigma y_{i-1}[f_1(G_i) - f_1(J_i)]| \leqq \mu \cdot \Sigma [f_1(G_i) - f_1(J_i)] < \frac{\mu}{n},$$

also

$$|\Sigma y_{i-1}[f_1(J_i) - f_1(A_i)]| < \frac{2\mu}{n}.$$

Für hinreichend großes n unterscheiden sich also die sämtlichen in (3) und (4) verglichenen Größen beliebig wenig von einander, womit die Gleichung (2) bewiesen ist.

Dies gilt natürlich für jede der vier Ableitungen. Also:

Ist $f(x)$ eine im Intervall (a, b) definierte Funktion mit beschränkten Differenzenquotienten, $g(x)$ eine ihrer vier Ableitungen, so ist

$$\int_a^b g(x) = f(b) - f(a).$$

Für irgend zwei dieser Ableitungen ist $\int_a^b (g_1(x) - g_2(x)) = 0$, und da dies auch für jedes Teilintervall gilt, so ist (S. 438) $g_1(x) = g_2(x)$ bis auf eine Menge vom Maße Null. Je zwei dieser Ableitungen und danach alle vier stimmen bis auf eine Menge vom Maße Null überein; bis auf eine Menge vom Maße Null hat also $f(x)$ eine einzige Ableitung $f'(x)$ oder jede Funktion mit beschränkten Differenzenquotienten ist bis auf eine Menge vom Maße Null differenzierbar.

Ist $f(x)$ in (a, b) integrabel, so hat ihr Integral

$$F(x) = \int_a^x f(x)$$

beschränkte Differenzenquotienten, denn $F(\alpha, \beta) = \int_\alpha^\beta f(x) : (\beta - \alpha)$ ist dem Betrage nach höchstens gleich der oberen Schranke von $|f(x)|$ im Gesamtintervall. Ist $g(x)$ eine der Ableitungen von $F(x)$, so ist

$$\int_a^b g(x) = F(b) = \int_a^b f(x),$$

und daraus folgt wie oben, daß bis auf eine Menge vom Maße Null $f(x)$ die Ableitung von $F(x)$ ist. Jede integrable oder meßbare (beschränkte) Funktion ist bis auf eine Menge vom Maße Null eine Ableitung, also Limes einer Folge stetiger Funktionen; man vergleiche dies mit den Betrachtungen über Bairesche Funktionen in § 4, IV.

Diese Sätze gewinnen natürlich noch an Interesse, wenn man sie von den durch Beschränkung auf eigentliche Integrale gebotenen Voraussetzungen befreit. Für den vorletzten Satz über die Differenzierbarkeit einer Funktion mit beschränkten Differenzenquotienten ist diese Verallgemeinerung ohne Herbeiziehung uneigentlicher Integrale möglich und gibt den schönen Satz von Lebesgue:

Jede stetige Funktion von beschränkter Variation ist bis auf eine Menge vom Maße Null differenzierbar.

Es genügt, dies für die monotonen Funktionen zu beweisen. Ist $y = f(x)$ im Intervall (x_0, x_1) stetig und monoton (nichtabnehmend) mit $y' \geqq y$ für $x' > x$, so ist, mit einer willkürlichen positiven Konstanten β,

$$t = x + \beta y$$

ebenfalls stetig und monoton (wachsend) mit $t' > t$ für $x' > x$. Führen wir t als unabhängige Variable ein, so werden $x = x(t)$, $y = y(t)$

Funktionen von t im Intervall (t_0, t_1) mit beschränkten Differenzenquotienten, denn es folgt für zusammengehörige Werte

$$t' - t = x' - x + \beta(y' - y),$$

$$0 \leqq \frac{x' - x}{t' - t} \leqq 1, \quad 0 \leqq \frac{y' - y}{t' - t} \leqq \frac{1}{\beta}.$$

Bis auf eine Menge T_0 (von Zahlen t) vom Maße Null sind also $x(t)$, $y(t)$ gleichzeitig differenzierbar. Ist ferner $\xi(t)$ die untere rechte Ableitung von $x(t)$, so ist

$$x_1 - x_0 = \int_{t_0}^{t_1} \xi(t),$$

und bedeutet T_1 die Menge der Stellen t, wo $\xi(t) = 0$, T_2 die der übrigen, also $(t_0, t_1) = T_1 + T_2$, so folgt, weil überall $\xi(t) \leqq 1$,

$$x_1 - x_0 \leqq f_1(T_2) = t_1 - t_0 - f_1(T_1),$$

also hat T_1 ein Längenmaß

$$f_1(T_1) \leqq t_1 - t_0 - (x_1 - x_0) = \beta(y_1 - y_0).$$

Bis auf die Menge $T = \mathfrak{S}(T_0, T_1)$ vom Maße $\leqq \beta(y_1 - y_0)$ sind also gleichzeitig die Ableitungen $\frac{dx}{dt} > 0$ und $\frac{dy}{dt}$, also auch die Ableitung $\frac{dy}{dx}$ vorhanden. Der Menge T entspricht als Bild vermöge $x = x(t)$ eine Menge X vom äußeren Maße $\varphi_1(X) \leqq f_1(T)$, wie man unmittelbar erkennt, wenn man T in ein lineares Gebiet einschließt und beachtet, daß einem t-Intervall stets ein x-Intervall von höchstens derselben Länge $x' - x \leqq t' - t$ entspricht. Die Menge der Stellen x, wo $\frac{dy}{dx}$ nicht existiert, ist $\leqq X$, hat also ein äußeres Maß $\leqq \beta(y_1 - y_0)$, also, da β beliebig ist, das Maß Null; die monotone Funktion $y = f(x)$ ist bis auf eine Menge vom Maße Null differenzierbar.

Anhang.

Nachträge und Anmerkungen.

Zusammenfassende Darstellungen der Mengenlehre sind:

A. Schoenflies, Die Entwickelung der Lehre von den Punktmannigfaltigkeiten. Erster Teil (Leipzig 1900), zweiter Teil (Leipzig 1908). Jahresbericht der Deutschen Mathematiker-Vereinigung, 8. Bd. und 2. Ergänzungsband.[1]

A. Schoenflies, Entwickelung der Mengenlehre und ihrer Anwendungen, erste Hälfte (Leipzig u. Berlin 1913).[2]

W. H. Young und Grace Chisholm Young, The theory of sets of points (Cambridge 1906).

Wir verweisen den Leser auf diese Werke insbesondere zur Ergänzung unserer Quellenangaben.

Erstes Kapitel.

§ 1.

S. 1. G. Cantor hat seine Forschungen in zahlreichen kleineren Abhandlungen niedergelegt, von denen ein Teil (in französischer Übersetzung) in Acta math. **2** (1883) gesammelt ist. Häufiger zu nennen sind die folgenden, als Punktmengen und Beiträge zitierten:

Über unendliche lineare Punktmannigfaltigkeiten:

I. Math. Ann. **15** (1879).
II. Math. Ann. **17** (1880).
III. Math. Ann. **20** (1882).
IV. Math. Ann. **21** (1883).
V. Math. Ann. **21** (1883).
VI. Math. Ann. **23** (1884).

[1] Als Bericht I, II zitiert. Mit Rücksicht auf den gedachten Leserkreis des vorliegenden Buches ist darauf hinzuweisen, daß der Schoenfliessche Bericht für Anfänger nicht geeignet ist, während der kritische Leser aus ihm reiche Anregung schöpfen und die erstaunlich umfassende Darstellung eines so ausgedehnten Gebietes nach Verdienst beurteilen wird.

[2] Diese Umarbeitung der entsprechenden Kapitel von Bericht I ist kurz vor beendigtem Satze des vorliegenden Buches erschienen und konnte nur noch in diesem Anhang berücksichtigt werden; sie ist als Mengenlehre zitiert.

V ist auch separat erschienen: Grundlagen einer allgemeinen Mannigfaltigkeitslehre (Leipzig 1883).

Beiträge zur Begründung der transfiniten Mengenlehre:

I. Math. Ann. **46** (1895).

II. Math. Ann. **49** (1897).

Der Ausdruck Paradoxieen des Unendlichen ist Titel einer Schrift von B. Bolzano (Leipzig 1851, 2. Aufl. Berlin 1889).

S. 2. E. Zermelo, Untersuchungen über die Grundlagen der Mengenlehre I, Math. Ann. **65** (1908). Die wirklichen und angeblichen Antinomieen der Mengenlehre haben zu ausgedehnten Kontroversen Anlaß gegeben. Das Paradoxon der „Menge aller Ordnungszahlen" wurde von C. Burali-Forti bemerkt: Una questione sui numeri transfiniti, Rend. Palermo **11** (1897); das mehr logischen als mathematischen Typus tragende Paradoxon der „Menge aller sich selbst nicht enthaltenden Mengen" von B. Russell, The principles of mathematics I (Cambridge 1903), ch. X. Vgl. die Darstellung bei G. Hessenberg, Grundbegriffe der Mengenlehre (Göttingen 1906) und das soeben erschienene nachgelassene Werk von J. König, Neue Grundlagen der Logik, Arithmetik und Mengenlehre (Leipzig 1914).

Hierher gehört auch die vielumstrittene Frage, unter welchen Bedingungen ein mathematisches Objekt, etwa eine Zahl, eine Menge, eine Funktion als „definiert" anzusehen sei (die Frage nach der Definition einer „Definition"). Wir folgen der freien Auffassung Cantors (Punktmengen III) und verlangen nicht, daß die logische Disjunktion, ob ein Ding einer Menge angehört oder nicht, mit unseren aktuellen Mitteln wirklich entschieden werden könne. Eine reelle Zahl ist entweder algebraisch oder transzendent, und damit ist die Menge der transzendenten Zahlen „wohldefiniert", obgleich man zur Zeit der ebengenannten Cantorschen Abhandlung noch nicht wußte, daß π zu ihr gehört (die berühmte Arbeit von F. Lindemann steht übrigens in demselben Annalenband an späterer Stelle), und heute noch nicht weiß, ob etwa π^π zu ihr gehört. Die Funktion $f(x)$, die für rationales x gleich 1 und für irrationales gleich 0 ist, ist wohldefiniert, obgleich wir die Werte $f(2^e)$, $f(2^\pi)$, $f(\pi^\pi)$ usw. nicht kennen. Dieser Mengenbegriff und dieser (Dirichletsche) Funktionenbegriff bindet sich weder an „Kriterien, die nur eine endliche Anzahl von Versuchen erfordern", noch an „analytische Darstellungen" u. dgl. Bekanntlich ist die gegenteilige, allerengste Einschränkung der mathematischen Definition von L. Kronecker gefordert, aber von niemandem je ernstlich durchgeführt worden.

§ 6.

S. 14. Zum Inhaltsproblem. Die Ankündigung am Schlusse des Paragraphen bezog sich auf eine ursprünglich geplante Darstellung der Inhaltstheorie (Kap. X), die jetzt zwar durch eine andere, kürzere ersetzt worden ist, über die aber folgende Angaben ein gewisses Interesse zu verdienen scheinen.

Nehmen wir an, den Mengen A eines Ringes $\mathfrak{M}$ seien bereits „Inhalte" $f(A)$ zugeordnet, die das Symmetrieaxiom erfüllen (S. 403),[1] d. h. für

$$A_1 = \mathfrak{S}(X_1, X_2), \quad A_2 = \mathfrak{D}(X_1, X_2)$$

sei

$$f(X_1) + f(X_2) = f(A_1) + f(A_2).$$

Dann gilt allgemein: sind $X_1, X_2, \ldots, X_m$ Mengen des Ringes und $A_1, A_2, \ldots, A_m$ ihre symmetrischen Grundmengen, so ist

$$f(X_1) + f(X_2) + \cdots + f(X_m) = f(A_1) + f(A_2) + \cdots + f(A_m).$$

Man beweist dies durch den Schluß von m auf $m+1$. Nehmen wir eine $(m+1)^{\text{te}}$ Menge Y hinzu und bezeichnen die symmetrischen Grundmengen von $X_1, \ldots, X_m, Y$ mit $C_1, C_2, \ldots, C_{m+1}$, so ist nach (5), S. 12 für $i = 1, 2, \ldots, m+1$

$$C_i = \mathfrak{S}\left(A_i, \mathfrak{D}(A_{i-1}, Y)\right),$$

wobei $A_{m+1} = 0$ und unter A_0 eine die Mengen X_i, Y umfassende Menge, etwa ihre Summe zu verstehen ist. Schreiben wir $\mathfrak{D}(A_i, Y) = D_i$, so ist

$$\mathfrak{S}(A_i, D_{i-1}) = C_i, \quad \mathfrak{D}(A_i, D_{i-1}) = D_i,$$

letzteres, weil $A_i \leqq A_{i-1}$. Danach ist

$$f(C_i) + f(D_i) = f(A_i) + f(D_{i-1}),$$

also

$$\begin{aligned} & f(C_1) + f(C_2) + \cdots + f(C_{m+1}) \\ = {} & f(A_1) + f(A_2) + \cdots + f(A_{m+1}) + f(D_0) - f(D_{m+1}) \\ = {} & f(A_1) + f(A_2) + \cdots + f(A_m) + f(Y) \\ = {} & f(X_1) + f(X_2) + \cdots + f(X_m) + f(Y), \end{aligned}$$

weil $A_{m+1} = D_{m+1} = 0$, $D_0 = Y$. Der Satz gilt also für $m+1$ Mengen, wenn er für m Mengen gilt.

Hiernach läßt sich zeigen, daß es, und zwar nur auf eine einzige Weise, möglich ist, auch den Mengen des kleinsten Körpers $\mathfrak{K} = \mathfrak{M}_\varkappa$ Inhalte zuzuordnen. Diese Mengen lassen sich nach S. 16 in der Form

$$A = (Z_1 - Y_1) + (Z_2 - Y_2) + \cdots + (Z_m - Y_m)$$

[1] Die Nullmenge soll den Inhalt 0 haben, sodaß aus dem Symmetrieaxiom speziell das Summenaxiom $f(X_1 + X_2) = f(X_1) + f(X_2)$ folgt.

darstellen, wo Z_i und $Y_i \leqq Z_i$ Mengen aus $\mathfrak{M}$ sind. Wenn eine die Inhalte von $\mathfrak{M}$ unverändert lassende und das Summenaxiom erfüllende Inhaltsbestimmung für $\mathfrak{K}$ möglich sein soll, so muß A jedenfalls den Inhalt

$$f(A) = \Sigma[f(Z_i) - f(Y_i)]$$

bekommen; es ist aber zu zeigen, daß dieser nur von A und nicht von der gewählten Darstellung abhängt. Ist nun gleichzeitig

$$A = (W_1 - V_1) + (W_2 - V_2) + \dots + (W_n - V_n),$$

so hatten wir S. 14 gesehen, daß die Mengensysteme Z_i, V_k und Y_i, W_k dieselben symmetrischen Grundmengen haben; nach dem verallgemeinerten Symmetriegesetze ist daher

$$\Sigma f(Z_i) + \Sigma f(V_k) = \Sigma f(Y_i) + \Sigma f(W_k),$$

also in der Tat

$$\Sigma[f(Z_i) - f(Y_i)] = \Sigma[f(W_k) - f(V_k)].$$

Da gleichzeitig $f(A) + f(B) = f(A + B)$, so befolgen die Inhalte der Mengen von $\mathfrak{K}$ das Summenaxiom und, weil es sich um einen Körper handelt, damit auch das Symmetrieaxiom.

Um auch den weiteren Gang der ursprünglich beabsichtigten Inhaltsdarstellung noch anzudeuten, werde in bezug auf ein bereits mit Inhalten versehenes System $\mathfrak{M}$ jeder Menge A durch „Interpolation" ein äußerer Inhalt $\varphi(A)$ und ein innerer Inhalt $\psi(A)$ zugeordnet, nämlich $\varphi(A)$ als untere Schranke der $f(P)$ und $\psi(A)$ als obere Schranke der $f(Q)$, wo $P \geqq A$ und $Q \leqq A$ Mengen aus $\mathfrak{M}$ sind. Diese äußeren und inneren Inhalte genügen den Symmetrie- oder Summenformeln Kap. X, § 2, (1), (2), (3), jenachdem $\mathfrak{M}$ ein Ring oder ein Körper ist. Es sei dann $\mathfrak{F}_0$ der Ring der abgeschlossenen Polygone, $\mathfrak{G}_0$ der Ring der Polygongebiete, $\mathfrak{K}_0$ der kleinste Körper über $\mathfrak{F}_0$ oder $\mathfrak{G}_0$; $\mathfrak{F}$ der Ring der (beschränkten) abgeschlossenen Mengen, $\mathfrak{G}$ der Ring der (beschränkten) Gebiete, $\mathfrak{K}$ der kleinste Körper über $\mathfrak{F}$ oder $\mathfrak{G}$. Die elementargeometrisch definierte Inhaltsbestimmung von $\mathfrak{F}_0$ (oder $\mathfrak{G}_0$) läßt sich zu einer solchen von $\mathfrak{K}_0$ erweitern; die aus $\mathfrak{K}_0$ (oder $\mathfrak{F}_0$ oder $\mathfrak{G}_0$ selbst) durch Interpolation gewonnenen äußeren und inneren Inhalte $\varphi_0(A)$, $\psi_0(A)$ entsprechen dem Peano-Jordanschen Standpunkt. Die äußeren Inhalte $\varphi_0(F)$ abgeschlossener Mengen, oder die inneren Inhalte $\psi_0(G)$ von Gebieten erfüllen das Symmetrieaxiom, und die damit gewonnene Inhaltsbestimmung von $\mathfrak{F}$ oder $\mathfrak{G}$ läßt sich nun zu einer (in beiden Fällen übereinstimmenden) Inhaltsbestimmung von $\mathfrak{K}$ erweitern. Endlich sind dann die aus $\mathfrak{K}$ durch Interpolation hervorgehenden Zahlen $\varphi(A)$, $\psi(A)$ das äußere und innere Lebesguesche Maß von A, wobei $\mathfrak{K}$ für die Bestimmung von $\varphi(A)$ auch durch $\mathfrak{G}$, für die Bestimmung von $\psi(A)$ auch durch $\mathfrak{F}$ vertreten werden kann.

§ 8.

S. 18. Der Ausdruck „fast alle“ von G. Kowalewski, Einführung in die Infinitesimalrechnung (Leipzig 1908), S. 14.

§ 9.

S. 21. Die Mengen Lim sup X_i und Lim inf X_i treten natürlich implicite an vielen Stellen in der mengentheoretischen Literatur auf; wie ich nachträglich sehe, hat sie schon E. Borel besonders betrachtet und als ensemble limite complet, ensemble limite restreint bezeichnet: Leçons sur les fonctions de variables réelles (Paris 1905), S. 18.

Zweites Kapitel.

§ 1.

S. 32. Das geordnete Paar (a, b) ist die natürliche Grundlage der Funktionsbeziehung. Wenn man es vermeiden und nur mit Mengen operieren will, so kann man eine Funktion $b = f(a)$ zunächst nur dann definieren, wenn die von a, b durchlaufenen Mengen A, B kein gemeinsames Element haben, und erst durch Vermittlung zweier solcher Funktionen den allgemeinen Fall realisieren; so verfährt E. Zermelo, Untersuchungen über die Grundlagen der Mengenlehre I, Math. Ann. **65** (1908). Da aber schon bei der grundlegenden Relation „a ist Element von b“ ($a \varepsilon b$, Zermelo) die Reihenfolge von a, b wesentlich ist, so scheint mir die Ausschaltung des geordneten Paares illusorisch.

Drittes Kapitel.

§ 1.

S. 46. Diese Erklärungen der Mächtigkeit bei Cantor, Beiträge I, S. 481 und B. Russell, The principles of mathematics I (Cambridge 1903), S. 115. Cantor bezeichnet im Sinne jener doppelten Abstraktion die Mächtigkeit von M mit $\overline{\overline{M}}$, den Ordnungstypus mit $\overline{M}$. Die sogenannte Definition durch Abstraktion (wobei man auf Grund einer symmetrischen, transitiven Relation die Dinge in „Klassen“ einteilt und das den Individuen einer Klasse Gemeinsame als selbständiges Gedankending hypostasiert) ist in der Mathematik typisch, leidet aber häufig an einer gewissen Unklarheit.

§ 2.

S. 48. Wenn man wie Schoenflies (Bericht I, S. 15, Mengenlehre, S. 33) unter A', B' echte Teilmengen versteht ($A' < A$, $B' < B$), so muß man zwischen endlichen und unendlichen Mengen unter-

scheiden; bei jenen ist der Fall (1) ausgeschlossen und (4) gibt Äquivalenz, bei diesen umgekehrt (nachdem der Wohlordnungssatz bewiesen und alle Mengen als vergleichbar erkannt sind). Es ist dies nicht der einzige Grund, aus dem wir den Sprachgebrauch, zu den Teilmengen von A die Menge A selbst nicht mitzurechnen, für durchaus unzweckmäßig halten.

Über die Literatur zum Äquivalenzsatz vgl. Schoenflies, Mengenlehre, S. 34.

§ 4.

S. 57. J. König, Zum Kontinuumproblem, Math. Ann. **60** (1904).

§ 5.

S. 61. Cantor, Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen, Journ. f. Math. **77** (1874).

S. 62. Der Leser, der sich für weitere Kuriositäten des Unendlichen interessiert, sei z. B. auf J. Richard, Sur la philosophie des mathématiques (Paris 1903) verwiesen.

S. 64. Die Unabzählbarkeit des Kontinuums entdeckte Cantor schon 1874 (Zitat zu S. 61). Der Beweis des Textes stammt ebenfalls von ihm: Über eine elementare Frage der Mannigfaltigkeitslehre, Jahresber. d. Deutschen Math.-Vg. **1** (1892).

S. 65. Cantor, Ein Beitrag zur Mannigfaltigkeitslehre, Journ. f. Math. **84** (1877). Hier ist zuerst die Kontinuumhypothese in der Form ausgesprochen, daß jede unendliche lineare Punktmenge entweder abzählbar oder von der Mächtigkeit des Kontinuums sei. Ebenda ist die Entdeckung mitgeteilt, daß auch die Ebene und der mehrdimensionale Raum nur die Mächtigkeit des Kontinuums haben.

Viertes Kapitel.

Zur Theorie der geordneten Mengen vergleiche man Cantors Beiträge, ferner (insbesondere auch zu Kap. VI) F. Hausdorff, Untersuchungen über Ordnungstypen I—V, Berichte der Sächs. Ges. d. Wiss. **58** (1906), **59** (1907); Grundzüge einer Theorie der geordneten Mengen, Math. Ann. **65** (1908), sowie neuere Arbeiten von P. Mahlo in den ebengenannten Leipziger Berichten (seit 1909).

§ 1.

S. 70. Daß eine Menge stets geordnet werden kann, d. h. daß es Funktionen $f(p)$ oder Paarmengen P der bezeichneten Art gibt, erscheint hier noch fraglich; später (Kap. V, § 7) wird aber bewiesen, daß jede Menge sogar wohlgeordnet werden kann.

S. 79. Für $\omega + \omega$ schreibt Cantor anfänglich 2ω, in den Beiträgen $\omega 2$.

§ 5.

S. 90. R. Dedekind, Stetigkeit und irrationale Zahlen (Braunschweig 1872).

§ 7.

S. 97. Der Beweis, daß die Mächtigkeit von $T(\aleph_0) \geqq \aleph$, ist von Cantor; daß sie $\leqq \aleph$, von F. Bernstein; beide mitgeteilt in Bernsteins Dissertation, Untersuchungen aus der Mengenlehre (Halle 1901). Der Satz läßt sich ohne Schwierigkeit dahin verallgemeinern, daß die Typenklasse $T(\aleph_\alpha)$ die Mächtigkeit $2^{\aleph_\alpha}$ hat; die Zahlenklasse $Z(\aleph_\alpha)$ hat die Mächtigkeit $\aleph_{\alpha+1} \leqq 2^{\aleph_\alpha}$ (vgl. S. 122).

S. 99. Satz II von Cantor, Beiträge I, S. 504. Die Menge R_a der rationalen Zahlen $> a$ hat also stets den Typus η; eine ähnliche Abbildung von R_0 auf R_a läßt sich offenbar zu einer stetigen, umkehrbar eindeutigen Abbildung $y = f(x)$ der ganzen Halbgeraden $x > 0$ auf die Halbgerade $y > a$ vervollständigen, bei der y mit x zugleich rational oder irrational ist (wie es für rationales a die triviale Abbildung $y = a + x$ tut).

Ordnet man einer beliebigen Folge natürlicher Zahlen $a_1, a_2, a_3, \ldots$ einerseits den unendlichen dyadischen Bruch

$$x = \left(\frac{1}{2}\right)^{a_1} + \left(\frac{1}{2}\right)^{a_1+a_2} + \left(\frac{1}{2}\right)^{a_1+a_2+a_3} + \cdots,$$

andererseits den Kettenbruch

$$y = 1:(a_1+1) - 1:(a_2+1) - 1:(a_3+1) - \ldots$$

zu, so ist dadurch (mit Hinzufügung von $f(0) = 0$) eine im Intervall $(0, 1)$ stetige monotone Funktion $y = f(x)$ definiert, bei der einem dyadisch rationalen x ein rationales y, einem rationalen, aber nicht dyadischen x ein quadratisch irrationales y entspricht. Vgl. H. Minkowski, Zur Geometrie der Zahlen, Verh. d. Heidelberger Kongr. (Leipzig 1905).

Fünftes Kapitel.

Zur Theorie der wohlgeordneten Mengen vgl. außer Cantors Beiträgen insbesondere G. Hessenberg, Grundbegriffe der Mengenlehre (Göttingen 1906), die erste allgemein gehaltene Darstellung.

§ 3.

S. 114. Weiteres über Normalfunktionen bei O. Veblen, Continuous increasing functions of finite and transfinite ordinals, Trans. Amer. Math. Soc. **9** (1908). Ferner sei verwiesen auf E. Jacobsthal, Über den Aufbau der transfiniten Arithmetik, Math. Ann. **66** (1908); Zur Arithmetik der transfiniten Zahlen, Math. Ann. **67** (1909).

§ 5.

S. 127. Den ersten (umständlicheren) Beweis der Gleichung $\aleph_\alpha^2 = \aleph_\alpha$ gab G. Hessenberg, Grundbegriffe der Mengenlehre, S. 593; der auf dem Diagonalverfahren beruhende Beweis wurde zuerst von Ph. Jourdain publiziert, On the multiplication of Alephs, Math. Ann. **65** (1908).

§ 7.

E. Zermelo, Beweis, daß jede Menge wohlgeordnet werden kann, Math. Ann. **59** (1904); Neuer Beweis für die Möglichkeit einer Wohlordnung, Math. Ann. **65** (1908).

Sechstes Kapitel.

§ 2.

S. 146. Den Existenzbeweis für Mengen vorgeschriebener Spezies habe ich geführt in den Grundzügen einer Theorie der geordneten Mengen, Math. Ann. **65** (1908), § 21. Anzahl der Spezies und Geschlechter § 22.

§ 3.

S. 158. Jede beschränkte, perfekte, nirgendsdichte Menge reeller Zahlen hat, der Größe nach geordnet, den Typus 2^{ω^*}.

§ 10.

Hierzu vgl. F. Hausdorff, Die Graduierung nach dem Endverlauf, Abhandl. der Sächs. Ges. d. Wiss. **31** (1909); Untersuchungen über Ordnungstypen V (Über Pantachietypen), Berichte der Sächs. Ges. d. Wiss. **59** (1907), wo man auch Literaturangaben und eine Kritik der vielfach unzutreffenden Vorstellungen über das Graduierungsproblem findet.

§ 11.

S. 199. H. Hahn, Über die nichtarchimedischen Größensysteme, Ber. d. Wiener Ak. d. Wiss. **116** (1907). Vgl. auch die dort zitierten Arbeiten, insbesondere G. Veronese, Fondamenti di geometria (Padua 1891), deutsch von A. Schepp (Leipzig 1894).

Siebentes Kapitel.

§ 1.

S. 210. Auf den Limesbegriff stützt sich die Theorie von M. Fréchet, Sur quelques points du calcul fonctionnel, Thèse (Paris 1906) = Rend. Palermo **22** (1906). Die Grundzüge der hier

entwickelten Umgebungstheorie habe ich im Sommersemester 1912 in einer Vorlesung über Mengenlehre an der Universität Bonn vorgetragen.

S. 211. H. Weyl, Die Idee der Riemannschen Fläche (Leipzig u. Berlin 1913).

S. 214. Bei einer beliebigen geordneten Menge (die also auch ein erstes oder letztes Element haben kann) sind alle Strecken (auch Anfangs- und Endstrecken), die x enthalten, als Umgebungen von x aufzufassen. Bei diesen zweiseitigen Umgebungen würde übrigens die Ordnung etwas in den Hintergrund gedrängt werden. Aber auch die rechtsseitigen Umgebungen von x, durch $x \leqq y$ oder $x \leqq y < b$ definiert, wie die linksseitigen, durch $y \leqq x$ oder $a < y \leqq x$ definierten, erfüllen die Umgebungsaxiome. — Mit der Übertragung der Punktmengentheorie auf geordnete Mengen beschäftigen sich A. Haar und D. König, Über einfach geordnete Mengen, Journ. f. Math. **139** (1908); H. Hahn, Über einfach geordnete Mengen, Ber. d. Wiener Ak. d. Wiss. **122** (1913).

§ 3.

S. 229. Man kann in naheliegender Verallgemeinerung die Menge $D = \mathfrak{D}_P P_\alpha$ betrachten, wo P alle in A enthaltenen Mengen einer gewissen Art oder $Q = A - P$ alle in A enthaltenen Mengen der entsprechenden Art durchläuft. Z. B., mit Bezug auf spätere Textstellen: Q eine Menge F_σ, oder Q eine Menge erster Kategorie in bezug auf A. Unter gewissen Umständen ist $D = 0$ das hinreichende (jedenfalls notwendige) Kriterium dafür, daß A selbst eine Menge Q ist.

§ 4.

S. 230. Der Satz I findet sich angedeutet bei Cantor, Punktmengen II.

S. 231. Für den Borelschen Satz sind viele, zum Teil unbegreiflich komplizierte Beweise gegeben worden, vgl. Schoenflies, Mengenlehre, S. 234.

§ 5.

S. 234. Wir hoffen hier in die Limesbildungen etwas System gebracht zu haben; die abgeschlossenen Limites (zu denen aber noch in Kap. VIII, § 6 der metrische Limes hinzutritt) sind, mehr oder minder klar, schon vielfach behandelt worden, z. B. Schoenflies, Bericht II, S. 109, während mir zu den Limesgebieten nur ein Ansatz von C. Carathéodory bekannt ist: Untersuchungen über die konformen Abbildungen von festen und veränderlichen Gebieten, Math. Ann. **72** (1912), S. 124.

§ 6.

Die auf eine beliebige Teilmenge M des Raumes E bezüglichen Relativbegriffe spielen in diesem Buch eine systematische Rolle; Keime dazu, in der Regel unter unnötigen Voraussetzungen über M (z. B. daß M perfekt oder nirgendsdicht sei) finden sich auch anderwärts. Die relative Ableitung in bezug auf eine beliebige Menge M im euklidischen Raume bei Schoenflies, Mengenlehre, S. 262.

§ 7.

S. 244. Für den Zusammenhang einer Menge gibt Cantor, Punktmengen V, S. 575 die hier auf S. 300, Anm. erwähnte Definition, die sich mit der unserigen nicht deckt, da z. B. nach ihr die Menge der rationalen Zahlen zusammenhängend ist. Sodann definiert C. Jordan, Cours d'analyse I, 2. Aufl. (Paris 1893), S. 25 und unabhängig von ihm Schoenflies, Beiträge zur Theorie der Punktmengen I, Math. Ann. **58** (1902) eine abgeschlossene (perfekte) Menge als zusammenhängend, wenn sie sich nicht in zwei abgeschlossene Teilmengen spalten läßt; ferner nennt E. Study, Geometrie der Dynamen (Leipzig 1903), S. 248 eine Punktmenge ein Kontinuum, wenn je zwei Punkte einer abgeschlossenen zusammenhängenden Teilmenge von ihr angehören. Ein Kontinuum ist auch in unserem Sinne zusammenhängend, aber nicht umgekehrt; z. B. ist die aus den Punkten $x > 0$, $y = \sin\frac{1}{x}$ und dem Koordinatenanfang $x = y = 0$ bestehende Punktmenge B zusammenhängend, aber kein Kontinuum. Das stetige Bild eines Kontinuums braucht kein Kontinuum zu sein (die gegenteilige Behauptung bei Schoenflies, Bericht II, S. 153, trifft allerdings dann zu, wenn man nur mit beschränkten euklidischen Mengen operiert). Die aus den Punkten $x > 0$, $y = \frac{1}{x}\sin\frac{1}{x}$ und der Ordinatenachse $x = 0$ bestehende Punktmenge A ist abgeschlossen und zusammenhängend; ordnet man jedem Punkt (x, y) von ihr den Punkt (x, xy) als Bild zu, so wird die obengenannte Menge B stetiges Bild von A, ohne ein Kontinuum zu sein.

§ 8.

S. 254. Cantor, Punktmengen V, S. 590.

§ 9.

S. 259. Nimmt man statt der Zahlenfolge einen Zahlenkomplex oder eine reelle Funktion $f(x)$, so hätte man y als eine Zahl u_α, u_β, u_γ zu bezeichnen, wenn die Ungleichung $y > f(x)$ für kein x, resp. nur endlich viele, resp. höchstens abzählbar viele x erfüllt ist; als eine

Zahl v_α, v_β, v_γ, wenn diese Ungleichung für mindestens ein x, resp. unendlich viele, resp. unabzählbar viele x erfüllt ist. Von den durch die Zerlegungen $T = U_\alpha + V_\alpha = U_\beta + V_\beta = U_\gamma + V_\gamma$ bestimmten Zahlen ξ_α, ξ_β, ξ_γ spielt im allgemeinen nur ξ_α, die untere Schranke der Funktion, eine Rolle; sie ist zugleich die untere Schranke der von den verschiedenen Funktionswerten gebildeten Zahlenmenge. Die Zahl ξ_β wäre nach Analogie als $\lim\inf f(x)$ zu bezeichnen, hat aber nichts mit dem sonst in der Analysis gebräuchlichen $\lim\inf_{x=a} f(x)$ zu schaffen, bei dem die S. 395 (für lim sup) beschriebene Einteilung der reellen Zahlen zugrunde liegt.

Achtes Kapitel.

§ 3.

Kugeln mit rationalen Mittelpunkten und Radien oder Würfel mit rationalen Eckpunkten sind schon mehrfach (von F. Bernstein, L. E. J. Brouwer u. a.) zur Vereinfachung der Theorie euklidischer Punktmengen benutzt worden. Unsere Darstellung dürfte erkennen lassen, daß die kleinste Mächtigkeit einer im Raume dichten Menge, oder eines mit dem gegebenen gleichwertigen Systems von Umgebungen, von zentraler Bedeutung und die Quelle aller Mächtigkeitstheoreme ist.

S. 269. Satz II verbunden mit § 9, IV enthält den sogenannten Cantorschen Hauptsatz (Punktmengen VI, S. 471), daß eine abgeschlossene Menge A in einen höchstens abzählbaren und einen perfekten Bestandteil zerfällt ($A = A_s + A_k = A_s + A_\gamma$).

S. 270. Satz III von Cantor, Punktmengen III mit Hilfe der Gebietsvolumina bewiesen.

S. 271. Satz IV von F. Bernstein, Untersuchungen aus der Mengenlehre (Halle 1901), S. 44.

S. 272. Die Sätze V, VI gehen auf W. H. Young, F. Bernstein und E. Lindelöf zurück, vgl. Schoenflies, Mengenlehre, S. 236, 244.

S. 273. Der Satz IX gilt übrigens für jeden metrischen Raum, auch ohne die Voraussetzung einer abzählbaren dichten Teilmenge.

§ 4.

S. 275. E. Lindelöf, Remarques sur un théorème fondamental de la théorie des ensembles, Acta Math. **29** (1905).

G. Cantor, Über verschiedene Theoreme aus der Theorie der Punktmengen in einem n-fach ausgedehnten stetigen Raume G_n, Acta Math. **7** (1885).

S. 284. Reduzible Mengen. Die S. 281 erwähnte Darstellung von $A - A_l$ oder der Menge A selbst, falls sie reduzibel ist, als Differenzensumme abgeschlossener Mengen lautet

(1) $$A = \sum_\xi (M_{2\xi} - M_{2\xi+1}),$$

wo

$$M_{2\xi} = A_{\xi\alpha}, \quad M_{2\xi+1} = \psi(A_\xi)_\alpha;$$

dies geht aus der Formel

$$M - \varphi(M) = M_\alpha - \psi(M)_\alpha$$

hervor, aus der zugleich wegen $M_\alpha \geqq \psi(M)_\alpha \geqq \varphi(M)_\alpha$ folgt, daß die Mengen M_ξ absteigend wohlgeordnet sind, d. h. $M_\xi \geqq M_\eta$ für $\xi < \eta$. Umgekehrt ist aber auch jede Menge der Form (1) mit absteigend wohlgeordneten, abgeschlossenen M_ξ reduzibel, wobei die Summe über alle Zahlen $\xi < \eta$ bei beliebigem η erstreckt werden kann: in Wahrheit werden ja nach einer höchstens abzählbaren Menge von Schritten die M_ξ gleich und die Summanden in (1) verschwinden. Wenn nämlich M abgeschlossen und A in $M - B$ abgeschlossen ist, so ist $\psi(A) = \mathfrak{D}(A_\alpha, B)$ in B abgeschlossen (vgl. S. 282 unten); wenn also M und M' abgeschlossen, $M \geqq M' \geqq C$ und A in $M - M' + C$ abgeschlossen ist, so ist $\psi(A)$ in $M' - C$ und $\varphi(A)$ in C abgeschlossen. Hiernach ist, wenn mit absteigenden, abgeschlossenen M_ξ die Gleichung (1) oder

$$A = (M_0 - M_1) + (M_2 - M_3) + (M_4 - M_5) + \dots$$

besteht, $\varphi(A)$ in $(M_2 - M_3) + (M_4 - M_5) + \dots$, $\varphi\varphi(A)$ in $(M_4 - M_5) + \dots$, also nach leichtem Induktionsbeweis jedes Residuum A_ξ in

$$(M_{2\xi} - M_{2\xi+1}) + \cdots$$

abgeschlossen, sodaß A_ξ schließlich verschwindet und A reduzibel ist. Die reduziblen Mengen sind also identisch mit den Differenzensummen absteigend wohlgeordneter, abgeschlossener Mengen; man kann diese Mengen als naturgemäße Übertragung der endlichen Differenzenketten mit assoziativem Gesetz (S. 9) ins Abzählbare ansehen. Während aber jede Summe von endlich vielen Differenzen abgeschlossener Mengen sich auch als endliche Differenzenkette mit absteigenden abgeschlossenen Mengen darstellen ließ, ist das gleiche bei unendlicher Gliederzahl nicht richtig. Z. B. ist die Menge R der rationalen Zahlen wohl als Summe von abzählbar vielen Differenzen abgeschlossener Mengen (mit Minuenden aus je einem Punkt und verschwindenden Subtrahenden) darstellbar, nicht aber mit absteigend wohlgeordneten abgeschlossenen Mengen, da ja eben R nicht reduzibel ist.

Der Leser möge zur Übung das Komplement $N - A$ in bezug auf eine abgeschlossene Menge N wieder als Differenzenkette darstellen: es ist

$$N - A = \sum_{\xi} (N_{2\xi} - N_{2\xi+1}),$$
$$N_0 = N, \quad N_\eta = \mathfrak{D}_{\xi} M_\xi \qquad (\xi < \eta),$$

wobei der Einfachheit wegen $M_0 \leqq N$ und die M_ξ schließlich verschwindend angenommen sind (die erste Annahme ist durch Schneiden von A mit N zu realisieren, die zweite evident zulässig).

Die reduziblen Mengen bilden einen Körper. Wir zeigen zunächst, daß der Durchschnitt $P = \mathfrak{D}(A, B)$ reduzibler Mengen wieder reduzibel ist. Das letzte Residuum P_l ist in P abgeschlossen, der Durchschnitt von P mit einer abgeschlossenen Menge F; $P_l = \mathfrak{D}(P, F)$ ist der Durchschnitt von $\mathfrak{D}(A, F)$ und $\mathfrak{D}(B, F)$, welche Mengen mit A und B ebenfalls reduzibel sind. Ändern wir die Bezeichnung dahin, daß wir die Durchschnitte von A, B, P mit F jetzt wieder A, B, P nennen, so haben wir zu zeigen, daß der Durchschnitt $P = \mathfrak{D}(A, B)$ reduzibler Mengen der Gleichung $\varphi(P) = P$ nur dann genügen kann, wenn er verschwindet.

Allgemein folgt aus $P = \mathfrak{D}(A, B)$ nach leichter Rechnung oder Überlegung: $\psi(P) \leqq \mathfrak{S}\big(\psi(A), \psi(B)\big)$.

Nehmen wir nun an, P verschwinde nicht, und setzen $\psi(P) = Q$, also

$$P_\alpha = P + Q = Q_\alpha.$$

Es ist dann $Q \leqq \mathfrak{S}\big(\psi(A), \psi(B)\big)$ oder

$$Q = \mathfrak{S}(A', B'), \quad A' = \mathfrak{D}\big(\psi(A), Q\big), \quad B' = \mathfrak{D}\big(\psi(B), Q\big).$$

Wir unterscheiden zwei Fälle:

(α) A' und B' seien beide in Q dicht.

Wir behaupten, daß dann jedes Residuum A_ξ von A die Teilmenge P enthält: $A_\xi \geqq P$. Für $A_0 = A$ ist dies richtig und offenbar auch für jedes A_η mit Limeszahlindex, wenn es für alle $\xi < \eta$ richtig ist; der Induktionsbeweis kann sich also auf den Schluß von ξ auf $\xi + 1$ beschränken. Wenn $A_\xi \geqq P$, so ist $A_{\xi\alpha} \geqq P_\alpha \geqq Q \geqq A'$. Nun haben A_ξ und A' keinen Punkt gemein, denn $A_\xi \leqq A$, $A' \leqq \psi(A)$. Durch Tilgung von A_ξ aus $A_{\xi\alpha}$ wird also kein Punkt von A' mitgetilgt, d. h. es ist noch

$$A_{\xi\alpha} - A_\xi \geqq A', \quad \psi(A_\xi) \geqq A'.$$

Da nun A' in Q dicht ist, so folgt

$$\psi(A_\xi)_\alpha \geqq A'_\alpha = Q_\alpha \geqq P.$$

$\psi(A_\xi)$ hat mit P keinen Punkt gemein; denn $P \leqq A$, $\psi(A_\xi) \leqq \psi(A)$

(weil A_ξ in A abgeschlossen ist, ist $\psi(A_\xi)$ in $\psi(A)$ abgeschlossen). Nach derselben Schlußweise wie soeben ist also noch

$$\psi(A_\xi)_\alpha - \psi(A_\xi) \geqq P, \quad \varphi(A_\xi) = A_{\xi+1} \geqq P.$$

Hiermit ist der Induktionsbeweis geliefert; dann ist aber $A_\iota \geqq P$ und A wäre nicht reduzibel.

(β) Eine der Mengen A', B' ist nicht in Q dicht.

Sei etwa B' in Q nicht dicht; d. h. es gibt ein nichtverschwindendes Relativgebiet $Q' = \mathfrak{D}(Q, G)$, das keinen Punkt von B' enthält und demzufolge ganz zu A' gehört: $Q' \leqq A'$. Wir setzen auch $P' = \mathfrak{D}(P, G)$ und erinnern uns (S. 241) der Formeln

$$P'_\alpha \geqq \mathfrak{D}(P_\alpha, G), \quad Q'_\alpha \geqq \mathfrak{D}(Q_\alpha, G),$$

also $$P'_\alpha \geqq P' + Q', \quad Q'_\alpha \geqq P' + Q';$$

auch P' verschwindet nicht.

Wir behaupten jetzt, daß jedes Residuum A_ξ die Menge P' enthält: wieder ist nur der Schluß von ξ auf $\xi + 1$ erforderlich, der genau so wie zuvor verläuft. Aus $A_\xi \geqq P'$ folgt $A_{\xi\alpha} \geqq P_\alpha' \geqq Q'$, und da $A_\xi \leqq A$ und $Q' \leqq A' \leqq \psi(A)$ keinen Punkt gemein haben: $\psi(A_\xi) \geqq Q'$. Sodann $\psi(A_\xi)_\alpha \geqq Q_\alpha' \geqq P'$, und da $P' \leqq A$ und $\psi(A_\xi) \leqq \psi(A)$ keinen Punkt gemein haben, $\varphi(A_\xi) \geqq P'$. Auch hier würde also $A_\iota \geqq P'$ und A nicht reduzibel sein.

Da nun das Komplement einer reduziblen Menge wieder reduzibel ist, so ist neben dem Durchschnitt auch die Summe und die Differenz zweier reduzibler Mengen wieder reduzibel; die reduziblen Mengen bilden einen Körper.

Wir fügen gleich hier noch folgendes hinzu. In einem metrischen Raume mit abzählbarer dichter Teilmenge ist eine reduzible Menge sowohl ein F_σ als auch ein G_δ (S. 306). In einem vollständigen Raume mit abzählbarer dichter Teilmenge ist dies umkehrbar, d. h. eine Menge A, die gleichzeitig ein F_σ und G_δ ist, reduzibel, sodaß hier also der Körper der reduziblen Mengen einfach der Durchschnitt der Ringe $\mathfrak{F}_\sigma$, $\mathfrak{G}_\delta$ ist. In der Tat: sei $A_\iota = P$ das letzte Residuum, $Q = \psi(P)$, ferner

$$M = P_\alpha = P + Q = Q_\alpha;$$

diese Menge ist perfekt, da P insichdicht ist (S. 281), also, falls sie nicht verschwindet, in bezug auf sich selbst von zweiter Kategorie (S. 328). Andrerseits aber ist P in A abgeschlossen, also selbst gleichzeitig ein F_σ und G_δ, ebenso Q, und beide in M dicht; also müßte P, als F_σ mit dichtem Komplement, in bezug auf M von erster Kategorie sein, ebenso Q, und M in bezug auf sich selbst von erster Kategorie. Dieser Widerspruch zeigt, daß P nicht von Null verschieden sein kann.

§ 5.

S. 287. D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen IV, Gött. Nachr. **8** (1906).

S. 288. Die Definition $\overline{xy} = \frac{1}{n}$ subsumiert die Untersuchungen von R. Baire über Elementfolgen unter unsere Theorie; vgl. Schoenflies, Mengenlehre, S. 386.

S. 290. Nennt man zwei, nicht identisch verschwindende Funktionen $x(t)$, $y(t)$ orthogonal, wenn $\int_0^1 x(t)y(t)\,dt = 0$, so ist in diesem Falle

$$\widehat{xy}^2 \geqq \int_0^1 x(t)^2\,dt > 0,$$

x also nicht Häufungselement von Funktionen, die zu x orthogonal sind. Eine Menge paarweise orthogonaler Funktionen ist also isoliert, demnach höchstens abzählbar. Vgl. F. Riesz, Sur les ensembles de fonctions, Compt. r. **143** (1906), S. 738.

§ 6.

S. 293. Zur Definition von $\overline{AB}$ vgl. D. Pompéju, Sur la continuité des fonctions de variables complexes, Ann. Fac. Toulouse (2) **7** (1905).

S. 298. Zum Folgenden vgl. L. E. J. Brouwer, On the structure of perfect sets of points, Amsterdam Akad. Proc. **12** (1910).

§ 7.

S. 304. E. Borel, Leçons sur les fonctions de variables réelles (Paris 1905), S. 17. Schoenflies nennt nur die Mengen G_δ Borelsche Mengen (Mengenlehre, S. 350).

S. 306. Daß eine separierte Menge ein G_δ ist, dürfte zuerst von W. H. Young bemerkt sein, Theory of sets (Cambridge 1906), S. 67, Theorem 35a.

S. 308. E. Borel, Leçons sur la théorie des fonctions (Paris 1898), S. 63, 93.

Die Menge D der Stellen, wo die Reihe (α) nicht konvergiert, ist, falls A insichdicht ist, zu A dicht (im allgemeinen Falle zum Kern A_k dicht). Man wähle nämlich die wachsenden natürlichen Zahlen $p, q, r, \ldots$ und die positiven Zahlen $\delta_p, \delta_q, \delta_r, \ldots$ folgendermaßen:

$$
\begin{aligned}
&p \text{ beliebig;}\\
2\delta_p &< \gamma_p;\\
&q \text{ so, daß } |a_q - a_p| < \delta_p;\\
2\delta_q &< \gamma_q,\ \delta_p,\ |a_1 - a_q|;\\
&r \text{ so, daß } |a_r - a_q| < \delta_q;\\
2\delta_r &< \gamma_r,\ \delta_q,\ |a_2 - a_r| \quad \text{usw.}
\end{aligned}
$$

Dann ist

$$x = a_p + (a_q - a_p) + (a_r - a_q) + \ldots$$

konvergent, und zwar

$$|x - a_p| < \delta_p + \delta_q + \delta_r + \ldots < \delta_p(1 + \tfrac{1}{2} + \tfrac{1}{4} + \ldots) = 2\delta_p,$$

ebenso

$$
\begin{aligned}
|x - a_q| &< 2\delta_q < |a_1 - a_q|,\\
|x - a_r| &< 2\delta_r < |a_2 - a_r| \quad \text{usw.},
\end{aligned}
$$

wonach x von $a_1, a_2, \ldots$ verschieden ist; ferner ist für $n = p, q, r, \ldots$

$$\left|\frac{c_n}{x - a_n}\right| > \frac{\gamma_n}{2\delta_n} > 1,$$

also die Reihe (α) nicht konvergent. Da man p beliebig und δ_p beliebig klein wählen kann, so liegt in jeder Umgebung eines Punktes von A ein Punkt von D.

S. 310. Das erste Beispiel einer solchen Reihe (α) dürfte von H. Bruns stammen, Bemerkungen zur Theorie der allgemeinen Störungen, Astr. Nachr. **109** (1884): es ist die Reihe

$$\Sigma \frac{\alpha^p \beta^q}{qx - p} = \Sigma \frac{c_{pq}}{x - \dfrac{p}{q}}, \quad c_{pq} = \frac{\alpha^p \beta^q}{q},$$

wobei α, β zwei feste positive Zahlen < 1 sind und p, q alle natürlichen Zahlen durchlaufen. Die Menge A ist die der positiven rationalen Zahlen, wobei das Auftreten jeder solchen in unendlich vielen Reihengliedern nur eine unwesentliche formale Differenz gegenüber unserer Darstellung ist. Beschränken wir uns auf positive x, so ist die Menge D der Nichtkonvergenzstellen dicht, hat aber das Längenmaß Null, weil $\Sigma c_{pq}^{\frac{1}{2}}$ konvergiert; die Menge C der Konvergenzstellen umfaßt alle irrationalen algebraischen Zahlen, weil $\Sigma \frac{c_{pq}}{\vartheta^q}$ für geeignetes $\vartheta < 1$ (nämlich $\beta < \vartheta < 1$) konvergiert.

§ 8.

S. 315. G. Cantor, Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen, Math. Ann. **5** (1872); E. Heine, Die Elemente der Funktionenlehre, Journ. f. Math. **74** (1872). Ch. Méray, Nouveau précis d'analyse infinitésimale (Paris 1872).

§ 9.

S. 318. Cantor, Punktmengen VI, S. 488.

S. 319. W. H. Young, Theory of sets, S. 64, Theorem 31.

S. 320. Die nächste Verschärfung des Youngschen Satzes, die sich noch mit denselben Hilfsmitteln beweisen läßt, lautet:

In einem vollständigen Raume mit abzählbarer dichter Teilmenge ist eine Menge $G_{\delta\sigma\delta}$, wenn sie unabzählbar ist, von der Mächtigkeit des Kontinuums.

Die fragliche Menge sei

$$P = \mathfrak{D}(A, B, C, \ldots),$$

wobei

$$A = \mathfrak{S}(A_0, A_1, A_2, \ldots),\ B = \mathfrak{S}(B_0, B_1, B_2, \ldots), \ldots$$

$$A_i = \mathfrak{D}(A_{i1}, A_{i2}, \ldots),\ B_k = \mathfrak{D}(B_{k1}, B_{k2}, \ldots), \ldots$$

und die Mengen A_{in}, B_{kn}, ... Gebiete sind. Wir können die A_i mit dem Index aufsteigend annehmen ($A_0 \leqq A_1 \leqq A_2 \leqq \ldots$), indem wir andernfalls A_i durch $\mathfrak{S}(A_0, A_1, \ldots, A_i)$ ersetzen, welche Menge als Summe endlich vieler G_δ wieder ein G_δ ist. Gleiches gilt für die B_k, C_l, ... Wir setzen noch

$$P_i = \mathfrak{D}(A_i, B, C, D, \ldots),\ P_{ik} = \mathfrak{D}(A_i, B_k, C, D, \ldots), \ldots.$$

Ist nun $P = \mathfrak{S}_i P_i$ unabzählbar, so ist mindestens eine der Mengen P_i unabzählbar. Da $P_0 \leqq P_1 \leqq P_2 \leqq \ldots$ und $P = \mathfrak{S}(P_i, P_{i+1}, \ldots)$, so können wir mit eventueller Nummernänderung bereits P_0 als unabzählbar annehmen. Es seien x_1, x_2 zwei zu P_0 gehörige Verdichtungspunkte von P_0 (deren Existenz durch § 3, I gesichert ist). Wir umgeben sie mit abgeschlossenen Kugeln V_1, V_2, die keinen Punkt gemein haben und dem Gebiet $G_1 = A_{01} \geqq A_0 \geqq P_0$ angehören; U_1, U_2 seien die entsprechenden Kugelgebiete mit gleichem Radius. In U_1 liegen unabzählbar viele Punkte von $P_0 = \mathfrak{S}_k P_{0k}$, also unabzählbar viele von mindestens einem P_{0k}; ebenso liegen in U_2 unabzählbar viele Punkte von einem P_{0k}; wegen $P_{00} \leqq P_{01} \leqq P_{02} \leqq \ldots$ können wir wieder annehmen, daß bereits P_{00} unabzählbar viele Punkte mit U_1 und U_2 gemein habe. Nun seien (für $p = 1, 2$) x_{p1}, x_{p2} zwei zu $\mathfrak{D}(P_{00}, U_p)$ gehörige Verdichtungspunkte dieser Menge. Wir umgeben sie mit abgeschlossenen Kugeln V_{p1}, V_{p2}, die keinen Punkt gemein haben, in V_p und außerdem in dem Gebiet

$$G_2 = \mathfrak{D}(A_{02}, B_{01}, B_{02}) \geqq \mathfrak{D}(A_0, B_0) \geqq P_{00}$$

liegen. In dem zu V_{pq} ($p, q = 1, 2$) gehörigen Kugelinnern U_{pq} liegen unabzählbar viele Punkte von $P_{00} = \mathfrak{S}_l P_{00l}$, also von mindestens einem P_{00l}; abermals können wir annehmen, daß dies für alle vier

Kombinationen p, q bereits auf P_{000} zutreffe. Sind dann x_{pq1}, x_{pq2} zwei zu $\mathfrak{D}(P_{000}, U_{pq})$ gehörige Verdichtungspunkte dieser Menge, so umgeben wir sie mit abgeschlossenen Kugeln V_{pq1}, V_{pq2}, die keinen Punkt gemein haben, in V_{pq} und außerdem in dem Gebiet

$$G_3 = \mathfrak{D}(A_{03}, B_{03}, C_{01}, C_{02}, C_{03}) \geqq \mathfrak{D}(A_0, B_0, C_0) \geqq P_{000}$$

liegen, und fahren so fort. Für jede aus den Ziffern 1, 2 gebildete Folge $p, q, r, \ldots$ hat man dann

$$V_p \leqq G_1, \ V_{pq} \leqq G_2, \ V_{pqr} \leqq G_3, \ \ldots;$$

läßt man überdies die Radien der V nach 0 konvergieren, so ist die dyadische Menge $\mathfrak{D}(\Sigma V_p, \Sigma V_{pq}, \Sigma V_{pqr}, \ldots)$ Teilmenge des Durchschnitts

$$\mathfrak{D}(G_1, G_2, G_3, \ldots) = \mathfrak{D}(A_0, B_0, C_0, \ldots);$$

dieser Durchschnitt und die Menge

$$P = \mathfrak{D}(A, B, C, \ldots) \geqq \mathfrak{D}(A_0, B_0, C_0, \ldots)$$

ist also (mindestens und auch höchstens) von der Mächtigkeit des Kontinuums.

In einem Raume ohne abzählbare dichte Teilmenge versagt dieser Beweis, weil dann eine unabzählbare Menge keine Verdichtungspunkte zu haben braucht. So ist der S. 289 erwähnte Raum von $\aleph_1$ Dimensionen mit den Punkten

$$x = (x_0, x_1, \ldots, x_\omega, \ldots)$$

vollständig, was ebenso wie für den Hilbertschen Raum bewiesen wird (S. 317); die Menge der Punkte, die nur eine Eins und sonst lauter Nullen als Koordinaten haben, ist von der Mächtigkeit $\aleph_1$ und hat keinen Verdichtungspunkt, nicht einmal einen Häufungspunkt, da ihre Punkte paarweise die Entfernung $\sqrt{2}$ haben. Auf sie ist also das Beweisverfahren nicht anwendbar, obwohl alle sonstigen Voraussetzungen zutreffen (die Menge ist divergent, also abgeschlossen, folglich ein G_δ und erst recht ein $G_{\delta\sigma}$ oder $G_{\delta\sigma\delta}$); wäre der Satz auch in diesem Falle richtig, so wäre damit das Kontinuumproblem im Cantorschen Sinne gelöst.

In einem vollständigen Raume mit abzählbarer dichter Teilmenge sind demgemäß die Mengen $G_{\delta\sigma\delta}$ und offenbar auch noch die Mengen $G_{\delta\sigma\delta\sigma}$ entweder endlich, abzählbar oder von der Mächtigkeit des Kontinuums; das trifft also auf die Mengen G, G_δ, $G_{\delta\sigma}$, $G_{\delta\sigma\delta}$, $G_{\delta\sigma\delta\sigma}$ und F, F_σ, $F_{\sigma\delta}$, $F_{\sigma\delta\sigma}$ zu. Der Versuch scheint nicht aussichtslos, das gleiche für alle Borelschen Mengen zu beweisen, d. h. die aus den Gebieten oder abgeschlossenen Mengen durch Summen- und Durchschnittsbildung über Folgen hervorgehen.

S. 321. In der euklidischen Ebene ist die Menge der Punkte mit rationalem x ein F_σ (Summe einer Folge von Geraden), aber kein G_δ, da sonst ihr Durchschnitt mit einer Geraden $y = \text{const}$ ein lineares G_δ wäre. Die Menge der Punkte mit irrationalem x ist ein G_δ, aber kein F_σ. Die Menge der Punkte mit rationalem x und irrationalem y ist weder ein F_σ noch ein G_δ, wohl aber von der Form $\mathfrak{D}(F_\sigma, G_\delta)$, also sowohl ein $F_{\sigma\delta}$ wie ein $G_{\delta\sigma}$.

S. 323. Vgl. das Zitat zu S. 298.

S. 328. R. Baire, Sur les fonctions de variables réelles, Annali di Mat. (3) **3** (1899).

§ 11.

S. 343. Ohne die Voraussetzungen über F kann der Satz VI versagen. Die Summe der Rechtecke Fig. 6, S. 249 (ohne die beiden Geraden) trennt einen rechts davon gelegenen Punkt von einem links gelegenen, obwohl es kein einzelnes Rechteck tut.

S. 346. L. E. J. Brouwer, Zur Analysis situs, Math. Ann. **68** (1910), S. 427.

S. 354. C. Jordan, Cours d'Analyse I (2. Aufl., Paris 1893), S. 91. Von dem Jordanschen Satz sind, teilweise unter einschränkenden Annahmen, viele Beweise gegeben worden, vgl. Schoenflies, Bericht II, S. 168; Young, Theory of sets, S. 225. Der hier reproduzierte von L. E. J. Brouwer, Beweis des Jordanschen Kurvensatzes, Math. Ann. **69** (1910); methodisch steht ihm am nächsten der von O. Veblen, Theory of plane curves in non-metrical analysis situs, Trans. Amer. Math. Soc. **6** (1905).

Neuntes Kapitel.

§ 1.

S. 363. Daß die Menge der Maximalwerte höchstens abzählbar ist, bemerkte Schoenflies, Bericht I, S. 157.

S. 368. Die Erweiterung einer in P gleichmäßig stetigen Funktion bei T. Brodén, Beiträge zur Theorie der stetigen Funktionen einer reellen Veränderlichen, Journ. f. Math. **118** (1897). Der Gegenstand ist vielfach behandelt worden, in besonders einfacher Weise von G. Faber, Über stetige Funktionen I, II, Math. Ann. **66** (1908) und **69** (1910).

§ 2.

S. 369. G. Peano, Sur une courbe qui remplit toute une aire plane, Math. Ann. **36** (1890). D. Hilbert, Über die stetige Abbildung einer Linie auf ein Flächenstück, Math. Ann. 38 (1891).

Die allgemeine Untersuchung der stetigen Kurven von Schoenflies, Bericht II, Kap. VI.

S. 373. Schoenflies, Bericht II, Kap. V.

S. 374. Die ersten Beispiele gaben W. F. Osgood, A Jordan curve of positive area, Trans. Amer. Math. Soc. **4** (1902), und H. Lebesgue, Sur le problème des aires, Bull. Soc. Math. **31** (1903). Nach F. Riesz, Sur les ensembles discontinus, Compt. r. **141** (1905), S. 650 kann man durch jede beschränkte punkthafte abgeschlossene Menge eine Jordansche Kurve legen.

S. 377. Zum Dimensionenproblem vgl. Schoenflies, Bericht II, S. 164.

S. 378. L. E. J. Brouwer, Beweis der Invarianz der Dimensionenzahl, Math. Ann. **70** (1911); Beweis der Invarianz des n-dimensionalen Gebiets, Beweis des Jordanschen Satzes für den n-dimensionalen Raum, Über Jordansche Mannigfaltigkeiten, Math. Ann. **71** (1912); Beweis der Invarianz der geschlossenen Kurve, Math. Ann. **72** (1912) u. a. Die den Leser zu vielen eigenen Ergänzungen zwingende Kürze der Brouwerschen Publikationen ist bei dem Mangel einer sonstigen einwandfreien und ausführlichen Darstellung sehr zu bedauern.

S. 379. E. Jürgens, Allgemeine Sätze über Systeme von zwei eindeutigen und stetigen reellen Funktionen von zwei reellen Veränderlichen (Leipzig 1879).

S. 381. Das Beispiel in der Anmerkung von E. Study, Vorlesungen über ausgewählte Gegenstände der Geometrie II (Konforme Abbildung einfach-zusammenhängender Bereiche) (Leipzig und Berlin 1913), S. 45.

§ 4.

S. 384. Wenn für jedes ε die Gleichung $\eta_n < \varepsilon$ für unendlich viele n und alle x erfüllt ist, nennt man die Folge einfach gleichmäßig konvergent; vgl. U. Dini, Grundlagen für eine Theorie der Funktionen einer veränderlichen reellen Größe (deutsch von J. Lüroth und A. Schepp, Leipzig 1892), S. 137. Wenn $f(x)$ mit keinem $f_n(x)$ identisch ist, ist einfach gleichmäßige und uniforme Konvergenz dasselbe.

S. 386. C. Arzelà, Sulle serie di funzioni di variabili reali, Rend. Bologna (2) **7** (1903).

S. 388. In IV sind Sätze und Beweise von Osgood, Baire, Van Vleck u. a. zusammengeflossen.

S. 389. R. Baire, Leçons sur les fonctions discontinues (Paris 1905); vgl. auch das Zitat zu S. 328. Die im Text nicht bewiesene Hälfte des Baireschen Satzes kommt darauf hinaus, daß für eine

Funktion $f(x)$, die nicht Limes stetiger Funktionen ist, eine in A abgeschlossene Menge P existiert, in der $f(x|P)$ überall unstetig ist. Dies hat besonders einfach H. Lebesgue bewiesen, Démonstration d'un théorème de M. Baire, Note II in E. Borel, Leçons sur les fonctions de variables réelles (Paris 1905).

Zehntes Kapitel.

§ 1.

S. 400. H. Hankel, Untersuchungen über die unendlich oft oszillierenden und unstetigen Funktionen, Math. Ann. **20** (1882). G. Cantor, Punktmengen VI, S. 473. G. Peano, Applicazioni geometriche del calcolo infinitesimale (Turin 1887), S. 153. C. Jordan, Remarques sur les intégrales définies, Journal de Math. (4) **8** (1892). E. Borel, Leçons sur la théorie des fonctions (Paris 1898), S. 46; Leçons sur les fonctions de variables réelles (Paris 1905), S. 16. H. Lebesgue, Intégrale, longueur, aire, Thèse (Paris 1902) = Annali di Mat. (3) **7** (1902); Leçons sur l'intégration (Paris 1904).

S. 402. Die Unlösbarkeit des Inhaltproblems. Wir zeigen, daß es nicht möglich ist, allen Punktmengen auf einer Kugelfläche K Inhalte zuzuordnen, die den Forderungen (α). (γ) von S. 401 genügen und wobei $f(K)$ positiv ausfällt. Der Beweis beruht auf der merkwürdigen Tatsache, daß eine Kugelhälfte und ein Kugeldrittel kongruent sein können, oder daß (von einer abzählbaren Menge abgesehen) K in drei Mengen A, B, C gespalten werden kann, die sowohl untereinander als auch mit $B+C$ kongruent sind. Nun müßte eine abzählbare Menge Q jedenfalls den Inhalt 0 haben; denn wählt man eine Drehungsachse, die durch keinen Punkt von Q geht, und einen Drehungswinkel, der keiner der geographischen Längendifferenzen zweier Punkte von Q gleich ist, so erhält man eine Drehung, die Q in eine Teilmenge von $P = K - Q$ überführt, und durch Wiederholung dieses Verfahrens erkennt man, daß K beliebig viele, paarweise fremde, mit Q kongruente Teilmengen hat, also $f(Q) \leqq \frac{1}{n} f(K)$ für $n = 1, 2, 3, \ldots$ und demnach $f(Q) = 0$ ist. Bei der angegebenen Zerlegung

$$K = Q + A + B + C$$

müßte also $f(A)$ gleichzeitig $= \frac{1}{2} f(K)$ und $= \frac{1}{3} f(K)$ sein, was der Annahme $f(K) > 0$ widerspricht.

Um eine solche Zerlegung herzustellen, verstehen wir unter φ eine Halbdrehung (um π) und unter ψ eine Dritteldrehung (um $\frac{2}{3}\pi$) um eine von der ersten verschiedene Achse. Sie erzeugen eine Gruppe G

von Drehungen, deren Elemente wir nach der Faktorenzahl geordnet (wobei φ, ψ, ψ^2 als einfache Faktoren gezählt werden) so schreiben:

(G) $$1 \mid \varphi, \psi, \psi^2 \mid \varphi\psi, \varphi\psi^2, \psi\varphi, \psi^2\varphi \mid \ldots.$$

Bei geeigneter Wahl der Drehungsachsen bestehen außer $\varphi^2 = \psi^3 = 1$ keine Relationen zwischen φ und ψ. Da dies der Nerv der ganzen Betrachtung ist, wollen wir es ausführlich beweisen. Die Produkte von zwei oder mehr Faktoren sind von einer der vier Formen

$$\alpha = \varphi\psi^{m_1}\varphi\psi^{m_2}\ldots\varphi\psi^{m_n}$$
$$\beta = \psi^{m_1}\varphi\psi^{m_2}\varphi\ldots\psi^{m_n}\varphi$$
$$\gamma = \varphi\psi^{m_1}\varphi\psi^{m_2}\ldots\varphi\psi^{m_n}\varphi$$
$$\delta = \psi^{m_1}\varphi\psi^{m_2}\ \ldots\varphi\psi^{m_n},$$

wo n eine natürliche Zahl und $m_1, m_2, \ldots, m_n$ gleich 1 oder 2 sind. Eine Relation $\varrho = \sigma$ zwischen zwei formal verschiedenen Produkten würde $\varrho\sigma^{-1} = 1$ zur Folge haben, d. h. ein von der Identität 1 formal verschiedenes Produkt müßte tatsächlich $= 1$ sein; dies Produkt müßte mindestens zwei Faktoren haben, und die fragliche Relation könnte in der Form $\alpha = 1$ angenommen werden (aus $\beta = 1$ würde $\varphi\beta\varphi = \alpha = 1$, aus $\gamma = 1$ ebenso $\varphi\gamma\varphi = \delta = 1$, aus $\delta = 1$ schließlich $\psi^{-m_1}\delta\psi^{m_1} = \alpha' = 1$ folgen). Es ist also zu zeigen, daß bei geeigneter Wahl der beiden Drehungsachsen alle Produkte α von 1 verschieden sind.

Legen wir durch den Kugelmittelpunkt ein rechtwinkliges Achsensystem, die ψ-Achse in die z-Achse, die φ-Achse in die xz-Ebene, bezeichnen den Winkel beider mit $\frac{1}{2}\vartheta$ und setzen

$$\lambda = \cos\frac{2}{3}\pi = -\frac{1}{2}, \quad \mu = \sin\frac{2}{3}\pi = \frac{\sqrt{3}}{2},$$

so lauten die unseren Drehungen entsprechenden orthogonalen Transformationen

$$(\psi) \quad \begin{cases} x' = x\lambda - y\mu \\ y' = x\mu + y\lambda \\ z' = z \end{cases}$$

$$(\varphi) \quad \begin{cases} x' = -x\cos\vartheta + z\sin\vartheta \\ y' = -y \\ z' = x\sin\vartheta + z\cos\vartheta \end{cases}$$

$$(\varphi\psi) \quad \begin{cases} x' = -x\lambda\cos\vartheta + y\mu + z\lambda\sin\vartheta \\ y' = -x\mu\cos\vartheta - y\lambda + z\mu\sin\vartheta \\ z' = x\sin\vartheta + z\cos\vartheta. \end{cases}$$

Durch Vertauschung von μ mit $-\mu$ tritt ψ^2 an Stelle von ψ. Nun bedeute α ein Produkt von n Doppelfaktoren $\varphi\psi$ oder $\varphi\psi^2$, $\alpha' = \alpha\varphi\psi$

oder $\alpha' = \alpha\varphi\psi^2$ ein Produkt von $n+1$ solchen; der Punkt mit den Koordinaten 0, 0, 1 gehe durch α in x, y, z, durch α' in x', y', z' über, sodaß zwischen diesen Koordinaten die Gleichungen $(\varphi\psi)$ oder die durch Vertauschung von μ mit $-\mu$ daraus hervorgehenden Gleichungen $(\varphi\psi^2)$ bestehen. Wir behaupten, daß $\frac{x}{\sin\vartheta}$, $\frac{y}{\sin\vartheta}$ Polynome $(n-1)^{\text{ten}}$ Grades, z ein Polynom n^{ten} Grades in $\cos\vartheta$ ist, also

$$x = \sin\vartheta\,(a\cos\vartheta^{n-1} + \ldots)$$
$$y = \sin\vartheta\,(b\cos\vartheta^{n-1} + \ldots)$$
$$z = c\cos\vartheta^{n} + \ldots.$$

Dies ist nämlich für $n = 1$ richtig, da der Punkt 0, 0, 1 durch $\varphi\psi$ oder $\varphi\psi^2$ in den Punkt $\lambda\sin\vartheta$, $\pm\mu\sin\vartheta$, $\cos\vartheta$ übergeht, und überträgt sich von n auf $n+1$, da nach den Gleichungen $(\varphi\psi)$ oder $(\varphi\psi^2)$

$$x' = \sin\vartheta\,(a'\cos\vartheta^{n} + \ldots)$$
$$y' = \sin\vartheta\,(b'\cos\vartheta^{n} + \ldots)$$
$$z' = c'\cos\vartheta^{n+1} + \ldots$$

ist. Für die höchsten Koeffizienten bestehen dabei die Formeln

$$a' = \lambda(c-a), \quad b' = \pm\mu(c-a), \quad c' = c-a,$$
$$c' - a' = (1-\lambda)(c-a) = \tfrac{3}{2}(c-a),$$

aus denen man durch wiederholte Anwendung $c - a = (\frac{3}{2})^n$ schließt. Die z-Koordinate des Punktes 0, 0, 1 wird also durch ein aus n Doppelfaktoren bestehendes Produkt α in

$$z = \left(\frac{3}{2}\right)^{n-1}\cos\vartheta^{n} + \ldots$$

transformiert und reduziert sich jedenfalls nicht identisch (für alle ϑ) auf den Wert 1. Es gibt also nur endlich viele Werte von $\cos\vartheta$, für die $\alpha = 1$ sein könnte, und es ist möglich, mit Vermeidung von höchstens abzählbar vielen Werten den Winkel ϑ so zu wählen, daß kein Produkt α gleich 1 wird.

Die hiernach nicht nur formal, sondern wirklich paarweise verschiedenen Drehungen (G) verteilen wir nunmehr auf drei Klassen A, B, C in der Weise, daß von zwei Drehungen ϱ, $\varrho\varphi$ die eine zu A, die andere zu $B + C$ und von drei Drehungen ϱ, $\varrho\psi$, $\varrho\psi^2$ je eine zu A, B, C gehört. Diese Verteilung ist möglich. Wenn sie nämlich für die Produkte von höchstens n Faktoren bereits geglückt ist (soweit die fraglichen Drehungen ϱ, $\varrho\varphi$ usw. höchstens n Faktoren haben), so läßt sie sich auch für die Produkte von höchstens $n+1$ Faktoren durchführen; denn ein Produkt von $n+1$ Faktoren ist entweder $= \varrho\varphi$, wo ϱ ein mit ψ oder ψ^2 endigendes

Produkt von n Faktoren ist, und werde, jenachdem ϱ zur Klasse A, B, C gehört, der Klasse B, A, A zugewiesen, oder es ist $= \sigma\psi$ oder $\sigma\psi^2$, wo σ ein mit φ endigendes Produkt von n Faktoren ist, und dann werde, jenachdem σ zu A, B, C gehört, $\sigma\psi$ der Klasse B, C, A und $\sigma\psi^2$ der Klasse C, A, B zugewiesen. Eine Kollision dieser Bestimmungen ist ausgeschlossen. Der Anfang des Verfahrens ist aus folgender Aufstellung ersichtlich (wenn wir die Identität 1 zur Klasse A rechnen):

A	1		$\psi\varphi, \psi^2\varphi, \varphi\psi^2$	$\varphi\psi\varphi$	...
B		φ, ψ		$\varphi\psi^2\varphi, \psi\varphi\psi, \psi^2\varphi\psi$	...
C		ψ^2	$\varphi\psi$	$\psi\varphi\psi^2, \psi^2\varphi\psi^2$	...

Endlich sei Q die abzählbare Menge der Fixpunkte (Drehungspole) der von 1 verschiedenen Drehungen unserer Gruppe und $K = P + Q$. Ein Punkt x von P geht durch die Drehungen (G) in die paarweise verschiedenen Punkte

$$x, x\varphi, x\psi, x\psi^2, \ldots$$

über, deren Menge P_x heiße, und zwei solche Mengen P_x, P_y sind entweder identisch oder haben keinen Punkt gemein. Aus jeder Menge P_x werde genau ein Punkt x ausgewählt und auf diese Weise die Menge $M = \{x, y, \ldots\}$ gebildet; dann ist

$$P = M + M\varphi + M\psi + M\psi^2 + \ldots.$$

Endlich zerfällt, gemäß der obigen Klasseneinteilung der Drehungen, P ebenfalls in drei Mengen, die wir wieder mit A, B, C bezeichnen wollen:

$$P = A + B + C,$$
$$A = M + M\psi\varphi + M\psi^2\varphi + M\varphi\psi^2 + \ldots$$
$$B = M\varphi + M\psi + \ldots$$
$$C = M\psi^2 + M\varphi\psi + \ldots$$

und nach Konstruktion ist

$$A\varphi = B + C, \quad A\psi = B, \quad A\psi^2 = C,$$

also die Mengen $A, B, C, B + C$ kongruent, womit der Beweis vollendet ist.

Eine den Forderungen (α), (β), (γ) von S. 401 genügende, alle beschränkten Mengen umfassende Inhaltsbestimmung ist also im drei- (und mehr-)dimensionalen euklidischen Raume unmöglich, da sie sonst auch auf der Kugel möglich wäre (indem man einer Menge auf der Kugel das Volumen des entsprechenden konischen Körpers als Inhalt zuordnet). Für die gerade Linie und die Ebene muß die Frage offen bleiben, da die Struktur der Bewegungsgruppe in diesen Fällen das obige Verfahren nicht zuläßt.

§ 4.

S. 418. Beispiele unmeßbarer Mengen bei Schoenflies, Mengenlehre, S. 374; das erste stammt von G. Vitali, Sul problema della misura dei gruppi di punti di una retta (Bologna 1905).

S. 419. E. Borel, Les probabilités dénombrables et leurs applications arithmétiques, Rend. Palermo **27** (1909).

S. 422. Das Kettenbruchproblem ist von H. Gyldén, A. Wiman, T. Brodén, E. Borel und F. Bernstein behandelt worden.

S. 429. M. Fréchet, Sur quelques points du calcul fonctionnel, Rend. Palermo **22** (1906). H. Lebesgue, Contribution à l'étude des correspondances de M. Zermelo, Bull. Soc. Math. **35** (1907).

§ 5.

Für das Folgende vgl. die zu S. 400 zitierten Werke von H. Lebesgue.

Register.

(Verweisung auf Seitenzahlen.)